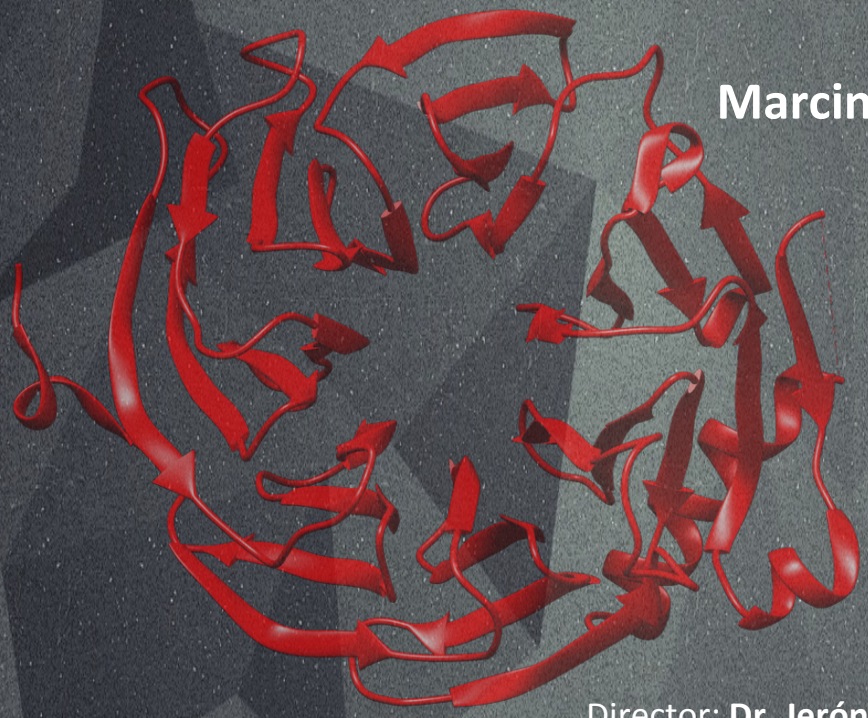
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Structural, biophysical and functional characterization of Nop7-Erb1-Ytm1 complex and its implications in eukaryotic ribosome biogenesis



Marcin Węgrecki

Director: **Dr. Jerónimo Bravo Sicilia**

Tutor: **Dr. María Adelaida García Gimeno**

September 2015

ibv INSTITUTO DE BIOMEDICINA DE VALENCIA CSIC

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

Structural, biophysical and functional characterization of Nop7-Erb1-Ytm1 complex and its implications in eukaryotic ribosome biogenesis

Marcin Węgrecki

A dissertation presented to the faculty of the Graduate School of the Polytechnic University of Valencia in candidacy for the degree of Doctor of Philosophy.

September 2015

Dr. **Jerónimo Bravo Sicilia**, supervisor of the Doctoral thesis developed by Mr. **Marcin Wegrecki** and titled **Structural, biophysical and functional characterization of Nop7/Erb1/Ytm1 complex and its implications in eukaryotic ribosome biogenesis,** which was carried out in Instituto de Biomedicina de Valencia

GIVES SUPPORT to the student's request of authorization for the deposit of the above-mentioned thesis.

Valencia, 15 of July of 2015

Signature:

Dr. **María Adelaida García Gimeno**, tutor of the Doctoral thesis developed by Mr. **Marcin Wegrecki** and titled **Structural, biophysical and functional characterization of Nop7/Erb1/Ytm1 complex and its implications in eukaryotic ribosome biogenesis**, which was carried out in Instituto de Biomedicina de Valencia

GIVES SUPPORT to the student's request of authorization for the deposit of the above-mentioned thesis.

Valencia, 15 of July of 2015

Signature:

To my mother.

To Ala, Ania and Marzenka

ACKNOWLEDGMENTS

After several years of hard work and scientific adventure it's time to stop for a moment and think about those who helped, who cared and who made it all possible. It is difficult to gather all the names together and it is possibly the most challenging task from all.

First of all, I would like to thank Dr Jerónimo Bravo, or Jero, my boss, for giving me the opportunity to taste science. The day we met for the interview was the beginning of a whole new era in my life. I became part of his lab and I started to learn a lot. I learnt how to fail and how to succeed. Thanks for including me in the UTS team. Thanks for giving me the chance to be part of our tiny family of locos.

Now it's time for those who are far away but who always believed in me. I want to thank to my family. Dziękuję Ci mamo za to, że nauczyłaś mnie jak być sobą, jak radzić sobie w trudnych momentach i jak nie być najgorszą osobą na świecie. Za to, że zawsze szanowałaś moje wybory i nigdy nie kwestionowałaś moich decyzji. Bez Twoich pierogów, mamo, doktorat byłby nie do zniesienia. Podziękowania należą się też mojej kochanej Cioci Basi. Ciociu, czasami trudno jest być w życiu dobrym ale Tobie ta sztuka wychodzi zawsze! Dzięki za to, że pomogłaś mi dążyć do celu, za to, że zawsze wierzyłaś i za wakacyjne wypady na Śnieżnik! Wielkie dzięki dla moich najlepszych na świecie sióstr! Marzenko, Ty zawsze wiedziałaś lepiej, zawsze podnosiłaś mnie na duchu i zawsze Ci zależało najbardziej. Dzięki Tobie i Andrzejowi udało mi się wiele osiągnąć i dzięki Wam za to i za super siostrzeńca. Dzięki Alu za bycie tą surową starszą siostrą. Dzięki Tobie nauczyłem się odrobiny porządku, a tego w mojej pracy najbradziej potrzeba. Dzięki za Wikę, nasze małe słoneczko. Najmłodsza ale równie ważna Aniu. Dzięki za bycie tą młodszą siostrą, bo to nie łatwe. Dzięki, bo wiem, że zawsze mogę na Ciebie liczyć i, że zawsze mnie odwiedzisz, gdziekolwiek bym nie wyjechał!

I would like to thank those that have struggled and laughed with me every day. Thanks Leti, Sara and Nada. We've become a family and I think that a great part of this work is thanks to you. I'm really grateful for all the support, all the help, all the motivation, ideas, good and bad moments. It is incredible how lucky I have been with you guys! I can't imagine anyone putting up with me as much as you did! Thanks Leti for the long hours of conversations and for many adventures that helped us to battle the stress. Thanks Nadita for the morning tostadas, for believing in me and for the support in the most difficult times. And Sarita, thank you for being my hermana valenciana, my first referee and my musical soulmate in the lab.

I thank Isa and Sandrea, former group members. You showed me how the science worked from the very first beginning and I have learnt a lot from your experience. I also want to acknowledge Alejandro, my former pre-grad student that did his best to learn and to carry out a part of this work.

Of course I owe an incredible gratitude to all those co-workers that have been around me during all these years and have influenced my research with useful tips and help. I would like to thank new members of the group, Susana and Jessica that have already showed me a lot of support. Thanks to Los Albertos for the seminars, for the ideas and criticism that were always very constructive. I thank all the colleagues from the institute because there was always someone to help, to offer a piece of advice and to teach me a new thing.

Thank you Ada for being my tutor, for spending your time and taking care of the bureaucracy. I really appreciate all the concern and words of support.

Thanks to all my friends that have been around me and always believed in me. Thank you Sara, my wonderful flat mate that has seen me for the four years of the PhD and patiently supported me. Thank you Stasio, it has been always a pleasure to have a good quality time that would hold the routine for a moment. Thanks Ola, Kasia and Kubon for our time together. Thanks my wonderful Polish team from Barcelona that has been always there. I know I have a second home with you guys. And Marta, my older friend scientist, thanks for all the craziness, love and support.

I need to thank my friends from the time when the idea of being a researcher was only an idea. Thank you Gema, Raquel, Soni and Bea for the time we shared in Salamanca. I know you always believed in me and you have always been there for me.

Last, but not least, I would like to acknowledge Jesús and Olga from Sevilla for teaching me the science a la andaluza. It was an incredible experience and I have learnt a lot. Thank you.

I know I should thank many more people. I would like to. I feel I owe many things to many persons that I have crossed my paths with but my memory has its limitations so one enormous thanks to everyone who has made this personal and professional project possible.

INDEX

LIST OF FIGURES AND TABLES

SUMMARY

Ribosome biogenesis is one of the most important and energy-consuming processes in the cell. However, the vast majority of the events and factors that are involved in the synthesis of ribosomal subunits are not well understood. Ribosome maturation comprises multiple steps of rRNA processing that require sequential association and dissociation of numerous assembly factors. These proteins establish a complex network of interactions that are essential for the pathway to continue. Extensive studies in Saccharomyces cerevisiae allowed to identify some of the genetic and functional correlations between the pre-ribosomal factors that could be organized into interdependent clusters or sub-complexes. A heterotrimer formed by Nop7, Erb1 and Ytm1 (PeBoW complex in mammals) is crucial for the proper formation of the 60S subunit. Depletion of any of the three proteins is inviable and certain truncations result in aberrant processing of 27SA2 rRNA thus impairing cell proliferation. Nop7 and Erb1 have been shown to bind RNA and are recruited to the pre60S before Ytm1. It is also known that the trimer has to be removed from the nascent particle in order to promote its normal maturation. Despite its relevance in the cell, the exact role of PeBoW is not clear and the interactions within the complex have been poorly characterized.

In this study we carry out an extensive biochemical and structural analysis of Nop7-Erb1-Ytm1 trimer from *S. cerevisiae* and from a thermophilic fungus *Chaetomium thermophilum*. We have been able to reconstitute a stable complex in vitro that was then used in crystallographic trials. We have solved the structure of the C-terminal domain of Erb1 from yeast that folds into a seven-bladed β-propeller. We prove that this part of the protein binds RNA in vitro, a property that might be important for its function. Moreover, in spite of previous reports suggesting that the β-propeller domain of Erb1 would not be essential for ribosome biogenesis, we could solve the crystal structure of Ytm1 bound to the carboxy-terminal portion of Erb1 from *C. thermophilum.* That finding led us to redefine the macromolecular interactions that hold the complex together. First, we have verified that the N-terminal region of Nop7 interacts with Erb1. Furthermore, we have shown that a good affinity binding takes place in vitro between WD40 domain of Ytm1 and the β-propeller of Erb1. Upon careful analysis of the interface involved in dimer formation we have designed a mutant of Erb1 that exhibits weaker association with Ytm1. We confirm our structural and biophysical data using *S. cerevisiae*. We prove that a point mutation that decreases the affinity between propellers of Erb1 and Ytm1 negatively affects growth in yeast because it interferes with 60S production. We show that a very conserved interface of protein-protein interaction could be targeted in order to hinder cell proliferation.

RESUMEN

El ensamblaje de ribosomas es uno de los procesos más importantes y costosos energéticamente en una célula eucariota. A pesar de ello, se sabe relativamente poco acerca de la gran mayoría de los eventos y factores implicados en la síntesis de las subunidades ribosomales. La maduración de ribosomas comprende numerosos pasos de procesamiento del rRNA que requieren la asociación y disociación de más de doscientos factores de ensamblaje. Esas proteínas establecen una compleja red de interacciones que son esenciales para que el proceso pueda llevarse a cabo. Los estudios realizados en *Saccharomyces cerevisiae* han permitido la identificación de algunas correlaciones genéticas y funcionales entre los factores prerribosomales. Es el caso del heterotrímero formado por Nop7, Erb1 e Ytm1 (complejo PeBoW en mamíferos), que es imprescindible para la correcta formación de la subunidad 60S. La ausencia de cualquiera de las tres proteínas es inviable y también se conocen ciertas variantes truncadas que alteran el procesamiento del rRNA $27SA_2$ y de este modo afectan la proliferación celular. Se ha demostrado que Nop7 y Erb1 se asocian al rRNA y que su reclutamiento al pre60S ocurre antes de la unión a Ytm1. Además se sabe que el trímero tiene que separarse de la partícula prerribosomal emergente con el fin de favorecer su maduración. A pesar de su gran relevancia en la célula, no está claro el papel exacto del complejo PeBoW y tampoco se dispone de conocimientos suficientes acerca de las interacciones intermoleculares que lo mantienen.

Durante el desarrollo de este proyecto se ha llevado a cabo un exhaustivo análisis bioquímico y estructural del trímero Nop7-Erb1-Ytm1 procedente de *S. cerevisiae* y del hongo termofílico *Chaetomium thermophilum.* En este trabajo hemos sido capaces de reconstituir el complejo estable *in vitro* que posteriormente se ha utilizado en los ensayos de cristalización, con los que hemos podido resolver la estructura del dominio carboxi-terminal de Erb1 de levadura, cuyo plegamiento corresponde a una hélice enrollada (*β-propeller*) de siete hojas. Gracias a la información estructural, hemos demostrado que esa parte de la proteína es capaz de unir RNA *in vitro*, lo que puede ser una propiedad importante para su función. Además, a pesar de los estudios anteriores que sugerían que la hélice enrollada de Erb1 no era esencial en la biogénesis del ribosoma, hemos resuelto la estructura cristalina de la proteína Ytm1 unida al dominio C-terminal de Erb1 de *C. thermophilum*. Ese descubrimiento nos ha permitido redefinir las interacciones macromoleculares que mantienen el complejo. Inicialmente hemos confirmado que el extremo amino-terminal de Nop7 interacciona con Erb1. A continuación, hemos demostrado que el dominio WD40 de Ytm1 se une al *β-propeller* de Erb1 con una buena afinidad. Después de un detallado análisis de la superficie involucrada en la formación del dímero, hemos sido capaces de diseñar una variante mutada de Erb1 que se asocia más débilmente con Ytm1. Los hallazgos estructurales y biofísicos se han confirmado *in vivo* usando *S. cerevisiae* donde hemos demostrado que una mutación puntual que disminuye la afinidad de unión entre los

dominios C-terminales de Erb1 e Ytm1 manifiesta un efecto negativo sobre el crecimiento de levadura porque interfiere con la síntesis de 60S. Nuestros resultados establecen un buen ejemplo de una superficie conservada involucrada en interacciones proteína-proteína, que podría considerarse una buena diana para inhibir la proliferación celular eucariota.

RESUM

L'ensamblatge de ribosomes és un dels processos més importants i energèticament costosos en una cèl·lula eucariota. Tot i això, es coneix relativament poc de la majoria dels factors implicats en la síntesi de les subunitats ribosomals. La maduració de ribosomes compren moltes etapes de processament del rRNA que requereix l'associació i dissociació de més de dos-cents factors d'ensamblatge. Aquestes proteïnes estableixen una complexa xarxa de interaccions que són essencials perquè el procés es pugi dur a terme. Els estudis realitzats en *Saccharomyces cerevisiae* han permès la identificació de algunes correlacions genètiques i funcionals entre els factors pre-ribosomals. Aquest és el cas del heterotrímer comprés per Nop7, Erb1 i Ytm1 (complex PeBoW en mamífers), que és imprescindible per a la correcta formació de la subunitat 60S. L'absència de qualsevol de les tres proteïnes és inviable i també és coneixen certes variants truncades que alteren el processament del rRNA $27SA_3$ i que d'aquesta manera afecten a la proliferació cel·lular. S'ha demostrat que Nop7 i Erb1 s'associen al rRNA i que el seu reclutament al pre60S té lloc abans de l'unió a Ytm1. A més a més, es sap que el trímer ha de separar-se de la partícula pre-ribosomal emergent per tal que es produeixi la seua maduració. Malgrat la seua rellevància en la cèl·lula, no s'ha aclarit el paper exacte del complex PeBoW i tampoc n'hi ha coneixements suficients de les interaccions intermoleculars que el mantenen.

Durant el desenvolupament d'aquest projecte s'ha dut a terme un exhaustiu anàlisi bioquímic i estructural del trímer Nop7-Erb1-Ytm1 de *S. cerevisiae* i del fong termofílic *Chaetomium thermophilum*. En aquest treball hem estat capaços de reconstituir el complex estable *in vitro* que posteriorment s'ha utilitzat en el assajos de cristal·lització, amb els que hem pogut resoldre l'estructura del domini carboxi-terminal de Erb1 de llevat i que té un plegament corresponent a una hèlix enrotllada (*β-propeller*) de set fulles. Gràcies a la informació estructural, hem pogut demostrar que aquesta part de la proteïna té la capacitat d'unir RNA *in vitro*, el que pot ser una propietat important per a la seua funció. A més a més, malgrat que els estudis anteriors suggerien que la hèlix enrotllada de Erb1 no era essencial en la biogènesis del ribosoma, hem pogut resoldre la estructura cristal·lina de la proteïna Ytm1 unida al domini C-terminal de Erb1 de *C. thermophilum.* Aquest descobriment ens ha permès redefinir les interaccions macromoleculars que mantenen el complex. Inicialment, hem confirmat que l'extrem amino-terminal de Nop7 interacciona amb Erb1. A continuació, hem demostrat que el domini WD40 de Ytm1 s'uneix al *β-*

propeller de Erb1 amb bona afinitat. Després d'un anàlisi detallat de la superfície involucrada en la formació del dímer, hem estat capaços de dissenyar una variant mutada de Erb1 que s'associa més dèbilment amb Ytm1. Les dades estructurals i biofísiques s'han confirmat *in vivo* utilitzant *S. cerevisiae* on hem demostrat que una mutació puntual que disminueix l'afinitat d'unió entre els dominis C-terminals de Erb1 i Ytm1 manifesta un efecte negatiu en el creixement del llevat perquè interfereix amb la síntesi del 60S. Els nostres resultats estableixen un bon exemple de una superfície conservada involucrada en interaccions proteïna-proteïna, que es podria considerar una bona diana per a inhibir la proliferació cel·lular eucariota.

1. INTRODUCTION

The central dogma of molecular biology (Fig. 1.1) stated by Crick in 1956 is a very concise description of the essential flow of information that takes place in all living organisms.

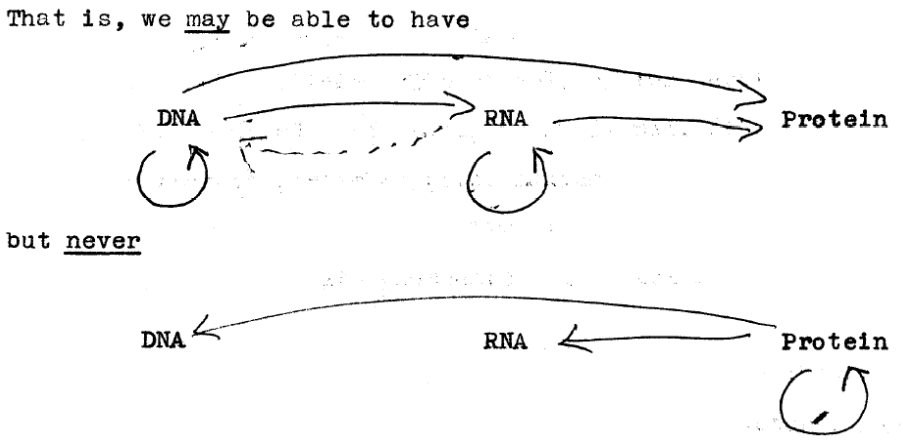


Figure 1.1. Central dogma of molecular biology.
A diagram representing the initial idea regarding the flow of information in living organisms. (From Crick, 1956)

Three types of macromolecules are involved in this most fundamental biological phenomenon (Berg et al., 2002). Deoxyribonucleic acid (DNA) is the carrier of the genetic code which can be copied during DNA replication. The functional information encoded in the sequence of DNA can be deciphered and transcribed into ribonucleic acid (RNA) that, on its own, can perform a certain range of functions. Nevertheless, the last step described by the dogma, that can evidence the full functional potential conserved in the genes, is the synthesis of proteins using mRNA as a template. This complex process called translation is carried out by ribosomes, large particles formed by rRNA and ribosomal proteins (r-proteins), that efficiently extract the information encrypted in the nucleic acid and convert it into polypeptides that finally will fold into functional proteins. Ribosomes are present in all organisms and are formed by two subunits: Small Ribosomal Subunit (SSU) that first binds to mRNA and controls the recognition of the genetic code by tRNAs and Large Subunit (LSU) that contains the peptidyl transferase center, where the peptide bond is formed (Melnikov et al., 2012; Ramakrishnan, 2002; Steitz, 2008). In

Prokaryotes the subunits are 30S (Svedberg, sedimentation rate) and 50S, respectively (Fig. 1.2). SSU 30S is made of 16S rRNA and 21 r-proteins, whereas 50S contains 5S rRNA, 23S rRNA and 31 r-proteins. More complex, Eukaryotic SSU, 40S, is formed by 18S rRNA and additional 33 r-proteins. 5S, 5.8S and 25S/28S rRNAs associate to 46 proteins to give a complete eukaryotic LSU, 60S. Fully functional ribosomes are one of the most complex macromolecular machines in the cell and constantly produce proteins that are essential for cell survival, growth and proliferation. The number and the activity of ribosomes is an important point of regulation of translation and gene expression in general (Barna, 2013; Deana, 2005).

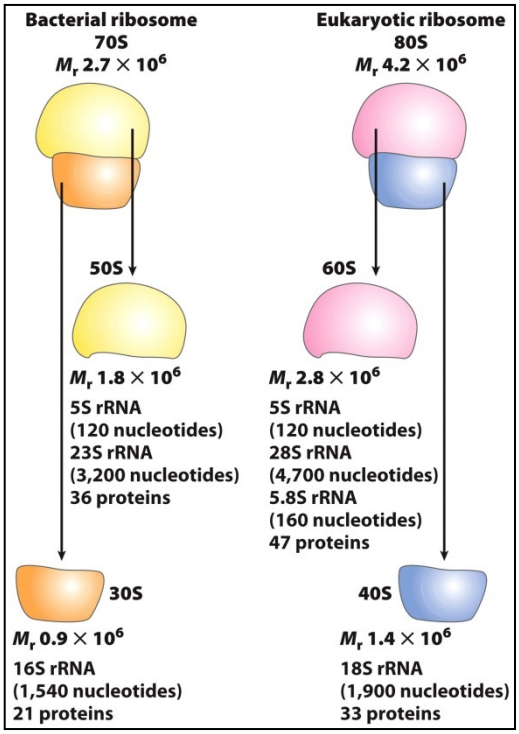


Figure 1.2. Comparison of prokaryotic and eukaryotic ribosomes.
Composition and molecular mass of the particles that form part of the ribosomes. (From Lehninger principles of biochemistry, 2008)

Thanks to the recent advances in structural biology we now have a clear picture of the architecture of procaryotic and eukarotic ribosomes (Ben-Shem et al., 2011; Ramakrishnan, 2002; Wimberly et al., 2000; Yusupov et al., 2001). It has been possible to crystallize and study in details the spatial organization of both subunits. X-ray crystallography has helped to explain the mechanism of translation,

its regulation, the basis of codon-anticodon recognition and finally the effect of diverse antibiotics that stall the translation by interfering with ribosome function (Jenner et al., 2013, 2010; Schmeing and Ramakrishnan, 2009). We have learned a lot about ribosomes in their final, mature and fully functional form but, in spite of their relevance in the cell, it is not yet fully understood how these large nucleoprotein macro complexes are correctly assembled during ribosome biogenesis, how is this process controlled and what are the factors that affect ribosome synthesis in the cell.

1.1. Eukaryotic ribosome biogenesis

The vast majority of studies regarding ribosome assembly have been done using *Saccharomyces cerevisiae* as a model organism thus this chapter mainly focuses on the process in yeast that, although generally conserved, presents some important differences when compared to higher Eukaryotes. The maturation of the ribosomal subunits is spatially divided between the nucleus and the cytoplasm. This compartmentalization introduces an additional step of regulation, thus indicating how tight is the control of the highly coordinated events that eventually will give rise to the mature 40S and 60S subunits. In general, production of the ribosomes is one of the most energetically expensive pathways in the cell (Warner, 1999). It requires a sequential participation of over 200 factors that allow proper rRNA processing and structural arrangement of the nascent subunits. Many of the trans-acting proteins are enzymes that provide activities such as GTPase, ATPase, exo- and endonuclease and helicase (Fig. 1.3) (Kressler et al., 2010; Matsuo et al., 2014; Rodríguez-Galán et al., 2013; Woolford and Baserga, 2013). Other factors constitute a structural scaffold that serves as a platform where the enzymes and the rRNA can interact in order to complete their functions. There are also described several types of snoRPNs that chemically modify nucleotides from rRNA or act as rRNA chaperones (Lafontaine, 2015). The association and dissociation of different factors is coupled to the shift of the pre-ribosomal particle from the nucleolus to the nucleoplasm and, finally, its export to the cytosol (Henras et al., 2008; Johnson et al., 2002; Nissan et al., 2002; Zemp and Kutay, 2007). The maturation of both subunits starts at the same time with the synthesis of a large precursor rRNA that will give rise to 5.8S, 18S and 25S/28S. In yeast, this firstly synthesized 35S rRNA is flanked by 5' and 3' External Transcribed Spacers (ETS) and contains two

Internal Transcribed Spacers: ITS1 that separates 18S from 5.8S and ITS2 that lies between 5.8S and 25S. The 5S rRNA is processed independently (Fernández-Pevida et al., 2015; Woolford and Baserga, 2013).

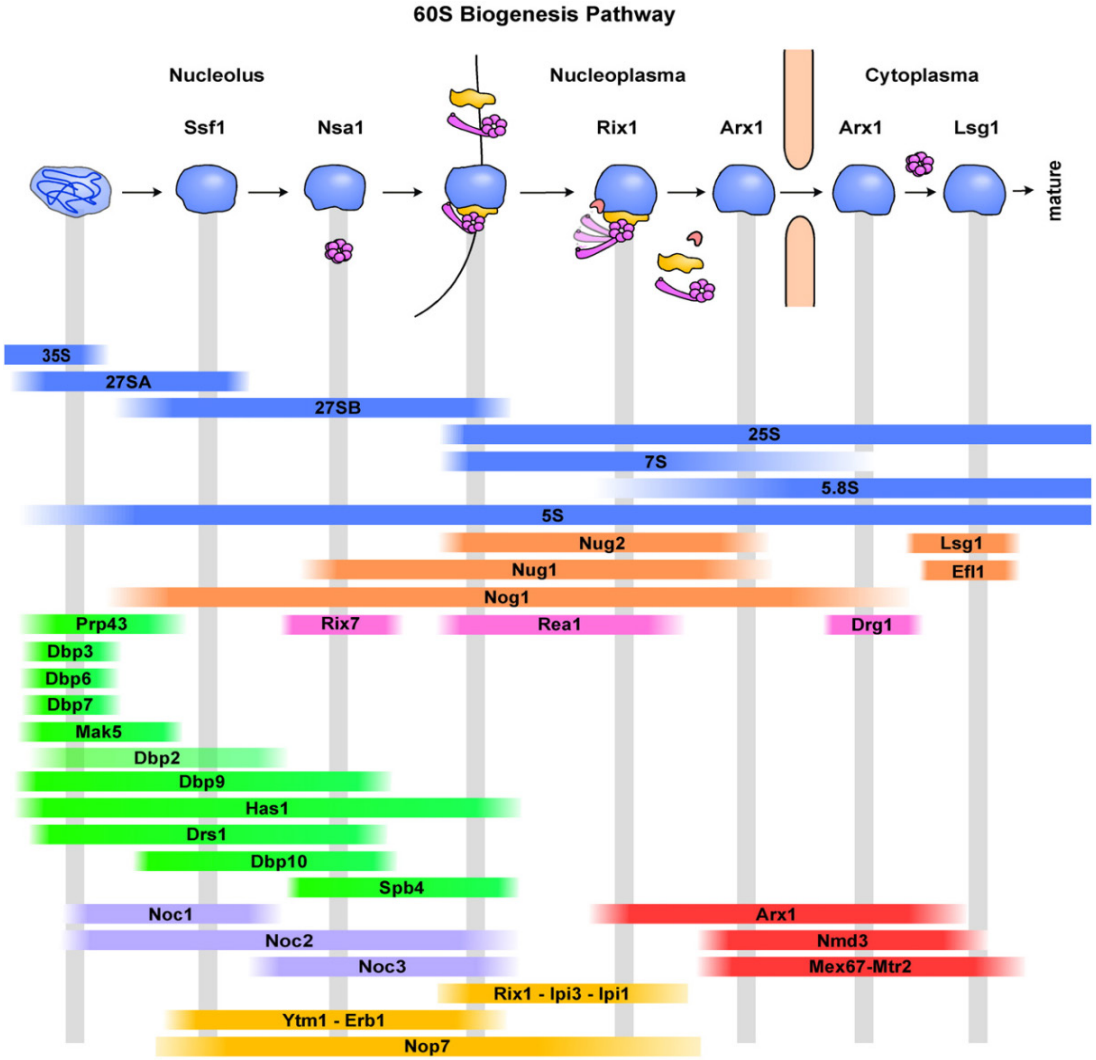


Figure 1.3. LSU assembly in Eukaryotes.
Sequential steps during biogenesis of 60S require participation of many assembly factors (AFs). rRNA species present at each step are depicted as blue lines, GTPases are shown in orange, ATPases in pink, helicases in green, export factors in red and important sub-complexes are depicted in violet and yellow. (From Kressler *et al*., 2010)

1.1.1. Transcription of rDNA and initiation of processing

The genes of ribosomal DNA (rDNA) are clustered in tandem repeats (approx. 9 kb long each) that, in yeast, encode all 4 rRNAs (Ganley and Kobayashi, 2007). As mentioned previously, 18S, 5.8S and 25S rRNAs are transcribed together, in the nucleolus, by RNA polymerase I (pol I) as one polycistronic chain that requires many steps of processing to give the mature, trimmed and modified rRNA particles (Fig. 1.4) (Moss et al., 2007).

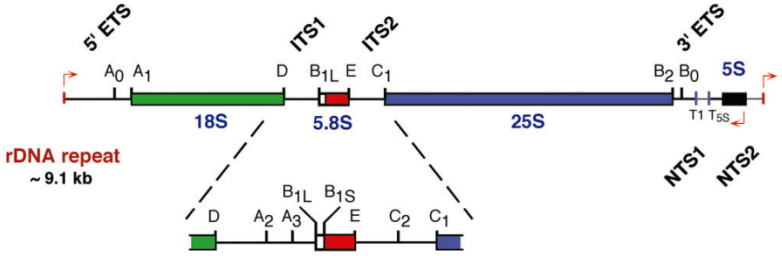


Figure 1.4 Organization of rDNA repeat in yeast.
A long transcript synthetized by pol I gives rise to 18S, 5.8S and 25S. 5S is transcribed separately by pol III. Red arrows indicate start of transcription. Removed spacers are marked (ETS, ITS). Cleavage sites that are targeted during rRNA processing are denoted with one-letter code. (From Fernández-Pevida *et al*., 2015)

5S rDNA is also present in the repeats but is transcribed separately, in the opposite direction, by RNA polymerase III. This is only true for yeast because in higher Eukaryotes 5S rDNA is located in different unrelated loci and its transcription does not occur in the nucleolus (Ciganda and Williams, 2011; Sørensen and Frederiksen, 1991). It has been questionable if 5S rRNA associates with pre60S particles early or rather in the final steps of ribosome maturation but recent studies seem to indicate that 5S rRNA joins 90S pre-ribosome in the initial stage of biogenesis. RNA pol I is the most active from the three Eukaryotic RNA polymerases because rRNA represents approximately 50% of all RNA synthesized in the cell (Albert et al., 2012; Russell and Zomerdijk, 2006). Before the transcription of polycistronic rDNA by pol I terminates, 5' end of the nascent 35S rRNA starts to associate with proteins and forms the first "processome" that will further develop into 90S pre-ribosomal particle (SSU processome) (Phipps et al., 2012). Although initially it was believed that the processing of rRNA would start

upon release of the transcript from pol I that is catalyzed by Rnt1 ribonuclease, recent findings have shown that in yeast first cleavage within ITS1 can occur co-transcriptionally thus quickly separating 40S and 60S maturation (Kos and Tollervey, 2010).

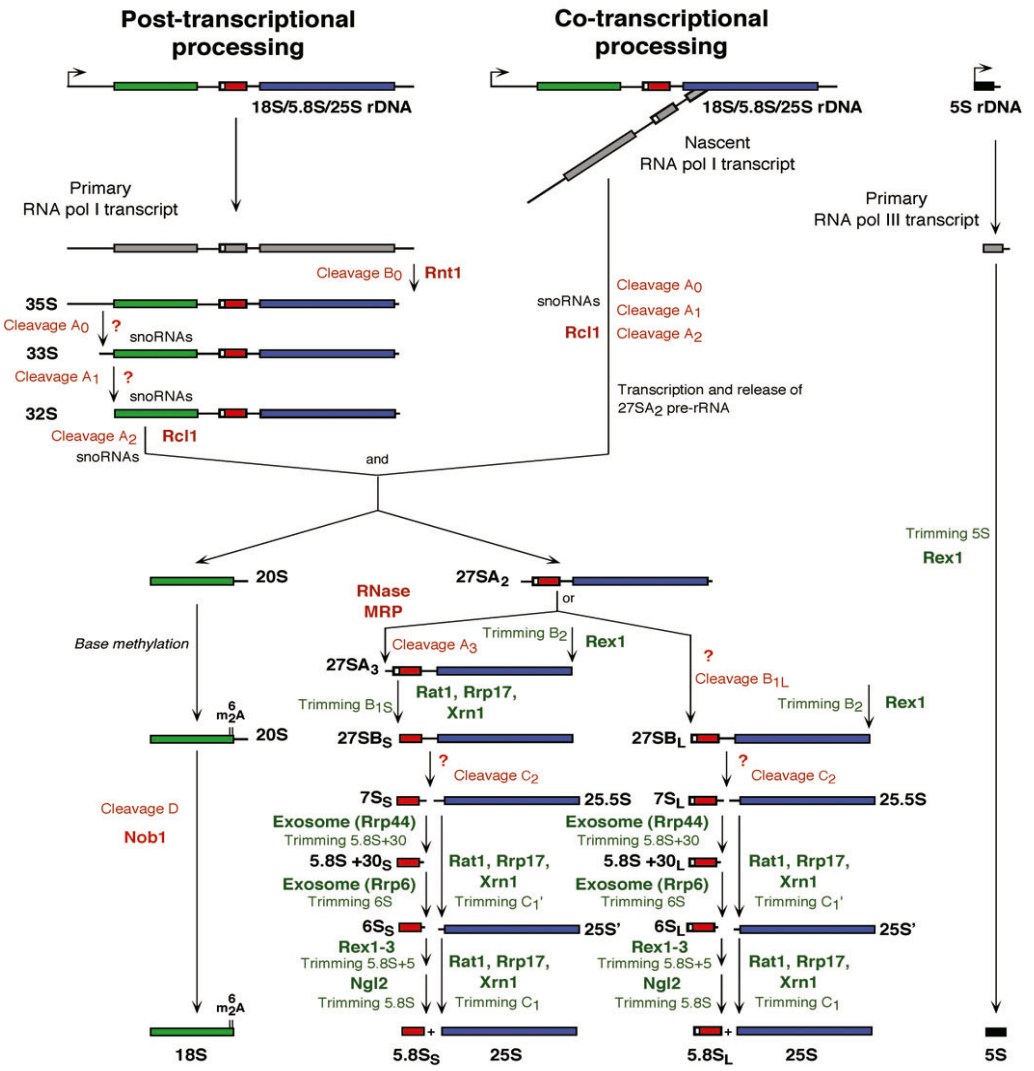


Figure 1.5 Processing of rRNA in yeast.
Detailed scheme of the nucleolytic events that produce mature forms of rRNA. Known enzymes involved in RNA cleavage are shown in the corresponding steps. For further details see text. (From Fernández-Pevida *et al*., 2015)

1.1.2. SSU rRNA processing

SSU contains mature 18S rRNA and 33 r-proteins. rRNA fragment corresponding to 18S particle is flanked by 5'ETS and ITS1 on its 3' end. Pre-18S is separated from pre-5.8S/23S part of the transcript by a cleavage at site A_2 within ITS1 that can occur either co-transcriptionally or post-transcriptionally (Fernández-Pevida et al., 2015) (Fig. 1.5). Rcl1 protein has been proposed to be involved in the process although the exact mechanism of A_2 cleavage is not yet clear (Horn et al., 2011). The removal of 5'ETS depends on the cleavage at A_0 and A_1 sites that require many factors including U3, snR30 and U14 snoRNAs and SSU processome (Lafontaine, 2015). The cleavages at A_0 and A_1 sites can occur almost simultaneously and in any order although the A_0 cleavage is dispensable. At the same time, cleavage at site A_1 has been functionally linked to that at A_2 within ITS1. In general these three events lead to formation of a 43S particle that contains yet immature pre18S rRNA. This pre40S complex is then exported to the cytoplasm where the last pre-ribosomal factors are released and rRNA processing is accomplished by Nob1 endonuclease that cleaves at site D and yields definitive 3' end (Lamanna and Karbstein, 2009). The last step of maturation seems to be LSU-dependent as it has been postulated that the Nob1-mediated cleavage is triggered within pre-40S/60S complex (Turowski et al., 2014).

1.1.3. LSU rRNA processing

Upon separation of pre-18S by cleavage at site A_2, resulting $27SA_2$ particle is further processed to form mature 25S and $5.8S_S$/$5.8S_L$ rRNAs. It is not clear how 3'-ETS is trimmed to give rise to the final 3' end of 25S but it has been suggested that 3'-ETS and ITS1 processing occur concomitantly (Fernández-Pevida et al., 2015). A recent study in mammals has proposed that DDX27 helicase (Drs1 in yeast) is indispensable for the 3' end maturation of 27S rRNA (Kellner et al., 2015). In eukaryotes there are known two pathways involved in 5.8S formation. A short 5.8S ($5.8S_S$) is generated by the major pathway whereas the minor pathway (15% of the processing) produces long 5.8S rRNA ($5.8S_L$) that retains 7-8 additional nucleotides on its 5' end (Fernández-Pevida et al., 2015; Henras et al., 2008; Woolford and Baserga, 2013). It is not yet understood what is the role of this heterogeneity of 5.8S but both are functional and found in mature 60S subunits (Henry et al., 1994). The

major pathway starts with a cleavage at A_3 site that is mediated by MRP endonuclease with a subsequent trimming done by Rat1 and Rrp17 that produces final $5.8S_S$ 5' end (Oeffinger et al., 2009). In the minor pathway, A_3 cleavage does not occur and unknown factor(s) perform endonucleolytic cut at site B_{1L} directly producing 5' end of the long 5.8S rRNA. Next, both particles $27SB_S$ and $27SB_L$ undergo the same processing that starts with the cleavage at site C_2 that splits 5.8S and 25S precursors (Fernández-Pevida et al., 2015). This step is complex and requires many pre-ribosomal factors but no known nuclease involved in the process has been identified up to date. Once separated, 25.5S is quickly shortened on its 5' end by Rat1 exonuclease and gives a mature 25S rRNA (Granneman et al., 2011). The processing of the 3' end of remaining $7S_S/7S_L$ rRNA is much more complex and requires a coordinated action of several nucleases. First, the exosome trims over 100 nt from the 3' end and then Rrp6 exonuclease continues until 6S particle, containing an overhang of 6-8 nt, is formed. At this stage, the pre-5.8S is exported to the cytoplasm where its 3' end is finally processed into the mature form by Ngl2 nuclease (Thomson and Tollervey, 2010).

1.1.4. Elements involved in 60S biogenesis.

In this work we will mainly focus on the formation of the LSU although the initial stages of the process, those that take place before first cleavage in ITS1, are common for both subunits. Given the complexity and relevance of the molecular events that occur during rRNA remodeling, it is not surprising that it requires the participation of numerous factors that coordinate and regulate ribosome biogenesis. Moreover, a sequential assembly of r-proteins with the nascent rRNA is indispensable to guarantee the particular architecture of the mature ribosome (Ohmayer et al., 2013).

1.1.4.1. r-proteins

Mature fungal ribosome contains 79 proteins, 64 of them are essential for growth (Steffen et al., 2012). The importance of r-proteins is affirmed by the fact that 59 of them possess a paralog copy encoded by a second gene (Korobeinikova et al., 2012). It also opens a possibility for ribosome specialization because it has been shown that the knock-out of one paralog does not always manifest the same

phenotype as the loss of function of its paralogous gene (Horiguchi et al., 2011; Schuldt, 2011; Wong et al., 2014; Yelick and Trainor, 2015). Moreover the expression of some of the paralogs has been correlated to specific tissues indicating again, that different forms of r-proteins may give rise to specific types of ribosomes (Xue et al., 2015). This is in line with the evolutive conservation of the minor pathway in 5.8S processing that provides additional variability. Ribosomal proteins play an important role in the assembly as they attach to rRNA modifying its structure, they serve as potential docking surfaces for binding of pre-ribosomal factors and in several cases they connect distant parts of the arising macro-complex through long extended motifs (Korobeinikova et al., 2012). Importantly, depletion of many r-proteins causes severe defects at different stages of ribosome biogenesis. Some of them associate with pre-rRNA early in the nucleolus whereas others join the subunit late, in the cytoplasm (Woolford and Baserga, 2013). Hierarchical attachment of ribosomal proteins to rRNA allows the creation of structural elements that will be present in the final 60S. It has been proposed that the position of each r-protein in the mature ribosome determines the step of biogenesis in which it might be involved (Woolford and Baserga, 2013). Several of the r-proteins have their pre-ribosomal counterparts that have similar sequence and structure but assemble earlier during ribosome biogenesis and are exchanged for the corresponding r-protein throughout the process. Those factors are so-called “place-holders” and occupy the position of specific r-proteins to prevent their premature association that could deregulate the pathway or produce aberrant, partially active complexes (Rodríguez-Mateos et al., 2009; Saveanu et al., 2003).

1.1.4.2. *trans*-acting factors

Proper rRNA processing, chemical modifications, association of r-proteins and spatial rearrangement of the forming ribosome strictly depend on an extremely complex and dynamic network of assembly factors (AFs) that, in many cases, are essential for growth and cell proliferation (Bassler et al., 2014; Gamalinda et al., 2014; Saveanu et al., 2003; Talkish et al., 2012). In general, during ribosome biogenesis, the composition of the earliest precursors is much more complex and the number of associated AFs gets reduced towards final steps of maturation. It is challenging to understand how all these factors work in an ordered and sequential

manner and how their action is controlled and linked to other cellular processes. From a functional point of view, the pre-ribosomal components involved in 60S synthesis can be classified into several groups:

a) snoRNPs: these small nucleolar ribonucleoproteins are formed by proteins and short RNAs that base-pair with compatible segments of nascent rRNA (Decatur et al., 2007; Henras et al., 2008; Lafontaine, 2015). Two different families of snoRNPs are responsible for nucleic acid modifications (Fig. 1.6). Box C/D snoRNA contains Nop1 methyltransferase (in yeast) and performs 2'-O-methylations of ribose. Box H/ACA is associated with Cbf5 that manifests pseudouridine synthase activity. Both types of modifications are found approximately 100 times in each ribosome and are believed to regulate its function. The processing of ITS1 at site A_3 is dependent on Mitochondrial RNA Processing (MRP) snoRNP that contains a short RNA and at least 10 proteins.

snoRNA	Class
U3	Box C/D
U14	Box C/D
snR30	Box H/ACA
snR10	Box H/ACA
RNase MRP	Unique

Figure 1.6 snoRNAs involved in RNA processing in yeast.
(From Woolford & Basegra, 2013)

b) nucleases: as already mentioned, a high number of exo- and endonculeolytic cleavages take place during rRNA processing (Fernández-Pevida et al., 2015; Henras et al., 2014; Woolford and Baserga, 2013). Although many of the enzymes involved in the process have been identified, there are still some missing elements of the nucleolytic machinery that trim rRNA (see sections 1.2 and 1.3 for more details)
c) NTPases: hydrolysis of ATP and GTP provides energy and can trigger a range of macromolecular events occurring during pre60S maturation (Kressler et al., 2012, 2008; Matsuo et al., 2014; Ulbrich et al., 2009). Several essential GTPases are

involved in different stages of the pathway. They induce conformational changes that allow dissociation and association of other proteins. Moreover, GTPase activity may also play role in the surveillance of the process marking check-points for important regulatory events. Three members of AAA-ATPase family have been shown to be crucial for the removal of assembly factors from pre-ribosomal particles. Rix7, Drg1 and Rea1 hydrolyze ATP and translate resulting energy into mechanical forces responsible for stripping off of several components during 60S biogenesis. Interesting mechanism of action has been proposed for Rea1 thanks to the structural information obtained by cryo-electron microscopy combined with biochemical analysis.

d) helicases: at least 10 RNA helicases participate in LSU maturation (Bernstein et al., 2006; Granneman et al., 2006; Kellner et al., 2015; Rodríguez-Galán et al., 2013). All of them, except one, belong to SF2 DExD/H-box family. Mrt4 is included into Ski2 family. In general, these enzymes present RNA-dependent ATPase activity and are able to unwind dsRNA, remodel ssRNA, facilitate nucleolytic cleavage of RNA chains and modulate RNA-protein associations. Interestingly some of these RNA-helicases are known to interact with AFs that can stimulate their ATP and/or unwinding activity.

e) functional clusters of non-enzymatic proteins: only a part of the assembly factors involved in LSU present enzymatic activities mentioned previously. Remaining pre-ribosomal proteins are believed to organize into sub-complexes that participate in different steps of rRNA processing and 60S maturation. Some of them act as, already described, place-holders that mimic r-proteins thus impeding their precocious association. Groups of AFs bind to pre-ribosomal particle and rearrange its architecture, stabilize rRNA, modulate the activity of the enzymes and perform quality control of the undergoing process. In many cases these proteins contain RNA-binding motifs or domains that are known to mediate high-affinity protein-protein interactions, like RRM or WD40 domains. This unambiguously indicates that a network of specific interactions must be maintained throughout the assembly pathway, probably due to its highly hierarchical organization. Several clusters or sub-complexes of functionally related proteins have been already associated with specific steps of ribosome biogenesis:

- Dbp6 helicase forms part of a complex with, at least, other 4 proteins (Npa1, Npa2, Nop8, Rsa3) and is indispensable for early steps of pre-60S maturation (de la Cruz et al., 2004).
- Nog1 GTPase has been shown to stably associate with Rlp24 protein that, in turn, binds Mak11 and these interactions are required for 27SB processing (Matsuo et al., 2014; Saveanu et al., 2007, 2003; Talkish et al., 2012). Actually, the three factors are members of a larger functional group of so-called B-factors that cooperatively lead to the loading of Nog2 GTPase on the nascent particle that will have decisive role in pre-60S export to the cytoplasm (Talkish et al., 2012).

1.1.4.3. Cluster of A_3-factors

Another well-described subset of factors is involved in processing of the $27SA_3$ rRNA and guarantee correct maturation of 5' end of $5.8S_S$ rRNA thus facilitating its association with 25S rRNA in the mature ribosome (Fig 1.7) (Granneman et al., 2011). The cluster of A_3-factors includes: Nop7, Erb1, Ytm1, Nop12, Cic1/Nsa3, Nop15, Rlp7 and Rrp1 and has been tightly related with the function of ribosomal proteins L7 and L8 (Harnpicharnchai et al., 2001; Jakovljevic et al., 2012; Mccann et al., 2015; Miles et al., 2005). Moreover, A_3-factors cooperate with exonucleases Rat1 and Rrp17, helicases Has1 and Drs1 in order to promote the trimming of the nascent $5.8S_S$ (Dembowski et al., 2013). While Rat1 binds to the early pre-60S particle in an independent manner, Rrp17 association is clearly conditioned by previous assembly of A_3 factors into the nascent complex (Granneman et al., 2011; Woolford and Baserga, 2013). In fact, although these AFs act within $27SA_3$ processing step, they join pre-ribosome much earlier. Surprisingly, when the binding of A_3 factors to rRNA was checked, it was discovered that they associated close to the area comprised by 5.8S 3'end – ITS2 – 25S 5'end (Ben-Shem et al., 2011; Granneman et al., 2011). This arose several hypotheses about the possible role of these 8 non-enzymatic proteins found in pre-60S processing. The binding of A_3-factors to ITS2 might prevent processing of this site before the complete removal of ITS1 occurs on the 5'end of $27SA_3$, a fact that, once again, could prove how ordered and sequential is the assembly pathway. At the same time it is known that A_3-factors must associate with pre-ribosome to stabilize the binding

of four r-proteins: L17, L26, L35 and L37 (Sahasranaman et al., 2011). It is then possible that A_3-cluster and r-proteins cooperatively remodel the architecture of $27SA_3$-containing particle thus promoting subsequent maturation events.

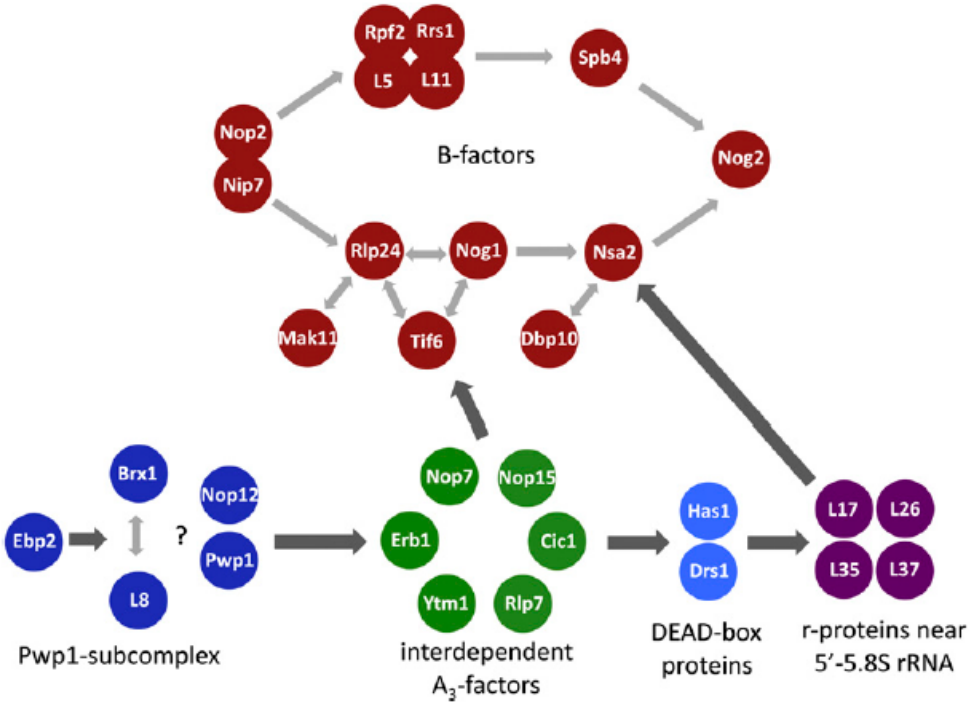


Figure 1.7 Functionally related clusters of proteins involved in early and middle steps of 60S assembly.
A complex network of functional interactions is established between numerous assembly factors. Pwp1 complex (blue) is required for a stable association of A_3-factors (green) that, in turn, condition the function of helicases (light blue) and binding of B-facotrs (red) and ribosomal proteins (purple) (From Woolford & Basegra, 2013)

It has also been shown that A_3 AFs should dissociate from the pre-ribosome in an interdependent manner in order to allow correct ribosome biogenesis. Null mutants for 7 members of A_3 cluster are inviable and only Nop12 protein is non-essential although it enhances 25S processing. It has been proposed that when A_3 factors work properly leading to a stable association of r-proteins, L17 binds close to the 5' end of $5.8S_S$ rRNA and physically blocks its further trimming by Rat1.

According to a model, when the A_3 cluster fails to carry out its function, L17 cannot stably bind and halt the exonucleolytic cleavage at B_{1S} site, therefore Rat1 continues to degrade 5' end of 27SB and initiate rRNA turnover (Sahasranaman et al., 2011). It is still to be elucidated how the proteins interact with each other during $27SA_3$ processing, what is the regulation of the process, which function corresponds to each member of the A_3 cluster. It was shown that within A_3-factors, three of them: Nop7, Erb1 and Ytm1 form a discrete heterotrimer that can be detected apart from the pre-ribosomal context (Miles et al., 2005; Tang et al., 2008). Nop7 sub-complex and its mammalian counterpart called PeBoW are essential in ribosome biogenesis but its exact role in the process is still under investigation.

1.2. Nop7 sub-complex

The trimer formed by Nop7, Erb1 and Ytm1 was firstly identified when Nop7 from yeast was studied (also known as Yph1p) (Du and Stillman, 2002). It was shown that it formed part of two complexes, while the larger one was formed by several r-proteins and a set of pre-ribosomal factors, the one corresponding to lower molecular weight contained, in addition to Nop7, only Erb1 and Ytm1.

1.2.1. Nop7

Nop7 (Nucleolar protein 7), called Pes1 (Pescadillo-homolog 1) in higher Eukaryotes, was firstly identified in zebra fish by *pescadillo* (*pes*) gene disruption that caused severe alterations in developing embryos and resulted in lethal phenotype (Auende et al., 1996). Further studies showed that the gene was highly conserved in all Eukaryotes and was essential for cell viability (Kinoshita et al., 2001; Sakumoto et al., 2001). Involvement of Pes1 in embryonic development has also been shown in *Xenopus laevis* (Tecza et al., 2011), whereas in plants it has been correlated to root morphogenesis (Zografidis et al., 2014). In *Candida albicans* it was described to be important in hypha-to-yeast switch and in yeast proliferation (Shen et al., 2008). Frequently the effect of *pes* disruption was tissue specific and the most striking phenotype was observed in areas of fast proliferation. In mammals, higher expression patterns have been detected in different types of cancer cells (Cheng et al., 2012; Kinoshita et al., 2001).

Finally, studies in *Saccharomyces cerevisiae* definitely associated Nop7 to ribosome biogenesis because a mutation within the gene had synthetically lethal phenotype together with a mutant Drs1 helicase (Adams et al., 2002). Subsequently it was co-purified with several pre-ribosomal complexes and its functional involvement into 5'end maturation of 5.8S was established (Oeffinger et al., 2002). Nop7 has been UV-crosslinked to domain III of 25S rRNA, close to L25 binding site. Tagged Nop7 co-precipitated 27S, 7S and 6S rRNAs indicating that its release from pre60S would occur late, after the cleavage at site C_2 from ITS2 (Granneman et al., 2011).

Nop7 has a BRCT domain in its central part which is believed to mediate protein-protein interactions and is required for nucleolar localization of the protein (Hölzel et al., 2007; Sakumoto et al., 2001). It also contains two predicted coiled-coils and a well conserved amino-terminal region (Pfam: 06732). Sequence inspection identified a nuclear localization sequence (NLS) and acidic regions towards the C-terminal part of the protein (Haque et al., 2000). Extensive mutational analysis of the human and the murine Pes1 have determined several segments which are essential for the function of the protein (Lapik et al., 2004). Interestingly, it was showed that mutant Pes1 would only exert a dominant negative phenotype if its association to Bop1 (Erb1 in yeast) was not abolished (Grimm et al., 2006; Hölzel et al., 2007).

Several binding partners of Pes1 have been described in mammalian cells including estrogen receptor (ERα and ERβ) (Cheng et al., 2012), Peter Pan (Ssf1 in yeast) (Fatica et al., 2002; Tecza et al., 2011) and B23 (Zhang et al., 2009). In a few studies, BRCT domain of Pes1 have been shown to interact with a protein-phosphatase Yvh1 and Pes1 overexpression could suppress slow-growing phenotype of Yvh1 loss-of-function (Sakumoto et al., 2001). Interestingly, it has been also reported that Nop7 sub-complex was found to be dissociated from pre-ribosomal particles when total cell extracts were treated with phosphatase inhibitors (Miles et al., 2005). The most recent reports on Nop7 function have suggested that it could act as a co-factor of Drs1 helicase because, at least under certain conditions, it was found to form part of the same small sub-complexes, tightly related to Nop7 (Merl et

al., 2010). Moreover, the human ortholog of Drs1, Ddx27 has been shown to interact with Pes1 in total cell extracts (Kellner et al., 2015).

From a structural point of view, Nop7/Pes1 is the only protein from the PeBoW complex with previously available information (Fig. 1.8). The structure of BRCT domain of human Pes1 in solution has been solved by Nuclear Magnetic Resonance (NMR) and deposited in Protein Data Bank with accession code 2EP8 (unpublished). It shows a β-sheet formed by four parallel β-strands with two short α-helices attached on one side of the sheet. This is a typical fold found in BRCT domain-containing proteins and superposes well (RMSD 1.9Å) with the BRCT domain of BRCA1 protein (PDB: 1T15) (Clapperton et al., 2004). Interestingly, BRCT domains have been predominantly found in proteins involved in DNA damage repair and cell cycle checkpoints (Gerloff et al., 2012). Pes1/Nop7, however, is not known to participate in any of these pathways.

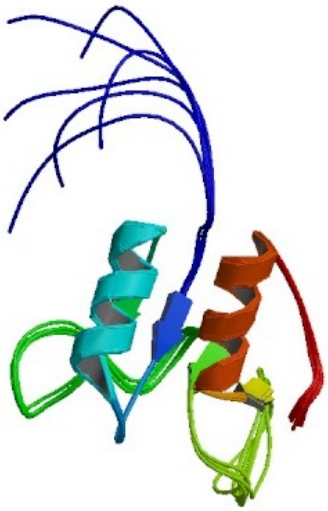


Figure 1.8 NMR structure of BRCT domain from human Pescadillo homolg 1.
The central domain of Pes1 is the only part of PeBoW complex with available structural information. (From PDB: 2EP8).

1.2.2. Erb1

Erb1, Eukaryotic ribosome biogenesis 1, (Bop1, Block of proliferation 1, in higher Eukaryotes) was firstly identified in mouse because an N-terminally truncated variant of the protein (Bop1Δ) induced a reversible growth arrest

(Strezoska et al., 2000). Further studies showed that this evolutionary conserved protein was predominantly found in the nucleolus and its possible role in ribosome biogenesis was investigated. Bop1Δ overexpression affected maturation of 28S rRNA (25S in yeast), led to the accumulation of the 36S precursor but did not show any effect on 18S processing (Pestov et al., 2001a; Strezoska et al., 2002). Those findings unambiguously indicated that Bop1 was involved in 60S biogenesis. *bop1* overexpression has been correlated with slower growth in mammalian cells and it increases the number of multipolar spindles during mitosis (Killian et al., 2004). Moreover, truncated Bop1 induced a reversible growth arrest through the p53 pathway thus it was proposed that Bop1 could link ribosome biogenesis to the cell cycle control (Pestov et al., 2001b). Deeper analysis of how Bop1 participated in rRNA processing was carried out upon identification of its locus in *Saccharomyces cerevisiae* (Harnpicharnchai et al., 2001; Horsey et al., 2004). Erb1 was quickly functionally correlated with other factors involved in $5.8S_S$ formation and was found to form part of pre-ribosomal complexes that also included Nop7, Rpl4, and Drs1, among others. Finally, Erb1 was shown to establish a tight physical interaction with Nop7 and Ytm1 proteins giving rise to Nop7 sub-complex formation (Lapik et al., 2004).

Initially, the sequence of Erb1/Bop1 indicated that it would harbor several WD40 repeats on its carboxy-terminal segment and that a highly conserved domain was present towards the N-terminal end (Pestov et al., 2001a). Moreover, a PEST sequence has been identified in the central region of the protein (Strezoska et al., 2000). PEST sequences are believed to appear in nuclear proteins that are tightly controlled by its fast turnover (Chevaillier, 1993). They contain a stretch of amino acids that is rich in proline, glutamic/aspartic acid, threonine and/or serine. These segments are quite flexible and might be easily accessible by proteases that can quickly degrade the protein. Interestingly, it has been shown that the stability of PeBoW complex depends on the stability of the individual factors so that the turnover of one of the proteins would affect the levels of the other two (Rohrmoser et al., 2007).

Similarly to Nop7, Erb1 has also been shown to UV-crosslink rRNA. In this case, it was found to bind helices H21 and H22 within the Domain I of 25S rRNA.

Co-precipitation experiments using HTP-tagged Erb1 demonstrated that it would efficiently link to 27S rRNA but the binding to 7S and 25S rRNA was weaker than the one observed for other A_3 factors (Granneman et al., 2011). This suggested that Erb1 would disassociate from the pre-ribosome right after C_2 cleavage but before Nop7 release. Since Nop7 and Erb1 directly interact on pre60S but their RNA binding sites are distant in the mature subunit it is believed that both proteins bring domains I and III of 25S rRNA together, thus allowing proper processing of the particle.

The most recent data have been pointing toward a possible association of Erb1 with Drs1 helicase (already known to be functionally linked to Nop7) and a large scale yeast two-hybrid analysis revealed that both proteins can interact (Mccann et al., 2015). Moreover, Ddx27 (mammalian counterpart of Drs1) has been co-precipitated with Bop1 in cellular extracts (Kellner et al., 2015). Nevertheless, it has not been show in vitro, without any possible secondary factors such as rRNA, that the binding between Drs1 and Erb1 or Nop7 takes place.

1.2.3. Ytm1

Ytm1 (Wdr12 in higher Eukaryotes) was first described in yeast as a microtubule binding protein that would affect chromosome stability when overexpressed (Ouspenski et al., 1999). Next, its murine counterpart, Wdr12, was identified as a novel WD-repeat protein found in the nucleus (Nal et al., 2002). Subsequent analysis showed that it contained a well conserved Notchless-like domain within its amino-terminal portion followed by 7 WD40 repeats that were predicted to fold into a seven-bladed β-propeller domain. Initially the function of Wdr12 was correlated to Notch signaling pathway and its association with Notch1 protein was tested. The sequence of the N-terminal domain of Wdr12 is similar to the amino-terminus of Notchless protein homolg 1 (Rsa4 in yeast) (Bassler et al., 2010). Recently, it has been shown that the N-terminal portion of Rsa4 is actually an Ubiquitin-like Domain (UBL) (Bassler et al., 2014). Eventually, studies on the pre-ribosomal complexes revealed that Ytm1 was found in Nop7-containing pre60S particles thus shedding light on its involvement in ribosome biogenesis. Polysome profiling upon Ytm1 depletion showed that 60S synthesis was heavily impaired (Miles et al., 2005). Moreover, Ytm1 remained bound to Nop7- and Erb1-containing low molecular

weight sub-complex even when ribosome biogenesis was altered in a temperature-sensitive mutant (*nop4-3*) (Harnpicharnchai et al., 2001).

Studies in the mammalian cells showed that a truncated Ytm1 lacking the UBL Domain (Wdr12ΔNle) was still found in the nucleolus and inhibited cell proliferation and rRNA processing in a dominant manner (Hölzel et al., 2005). Additional observations indicated that the disruption of the WD40 repeats from C-terminal domain induced a dispersed cytoplasmic and nucleoplasmic localization. Finally it has been confirmed that the mammalian Wdr12, likewise its fungal counterpart Ytm1, formed a stable complex with Pes1 and Bop1 (Nop7 and Erb1 in yeast) (Miles et al., 2005). Temperature-sensitive mutant *ytm1-1* altered the maturation of the pre-ribosome in yeast and decreased the association of Ytm1 with Erb1 and with pre60S. Ytm1-1 contained two mutations G398D and S442N, both within WD40 repeats of its C-terminal domain. That was the first indication that the β-propeller domain of Ytm1 mediated the interaction with Erb1 and was sufficient for the nucleolar localization of the protein (Tang et al., 2008).

Further studies in *Saccharomyces cerevisiae* provided an explanation for the role of the ubiquitin-like domain of Ytm1. Biochemical and functional analyses showed that a conserved glutamic acid present within the N-terminal segment was crucial for its interaction with MIDAS (Metal Ion Dependent Adhesion Site) domain of Rea1 (Bassler et al., 2010). Rea1 (Midasin in mammals) is a large AAA-ATPase that is essential in ribosome assembly because it transforms chemical energy into mechanochemical forces necessary for the removal of some pre-ribosomal factors (see section 1.4.2). A glutamic acid residue (E80 in *S.cerevisiae*) from UBL of Ytm1 (and Rsa4 in a subsequent step of ribosome biogenesis) coordinates a Mg^{2+} ion altogether with other 5 residues from MIDAS of Rea1 (Kressler et al., 2012; Ulbrich et al., 2009). The relevance of this interaction has been supported by the fact that Ytm1 carrying E80A mutation causes dominant-lethal phenotype by interfering with the 60S assembly. The same study demonstrated that Rea1 was involved in the release of Ytm1, Erb1 and (at least partially) Nop7 from the nascent ribosome and that the dissociation of the sub-complex was indispensable for the maturation to continue.

1.2.4. Nop7 sub-complex formation

Several studies were carried out to shed light on the interactions between Nop7, Erb1 and Ytm1 and their function in ribosome biogenesis (Hölzel et al., 2005; Miles et al., 2005; Rohrmoser et al., 2007; Tang et al., 2008). Erb1 is considered to be the core of the complex because it physically interacts with both, Nop7 and Ytm1. It has been proposed that the association of the proteins with pre-ribosomal particles was interdependent and a fully functional trimer had to be assembled in order to promote its stable binding to pre60S. Investigations in mammalian cells and yeast led to a conclusion that Nop7 and Erb1 joined the pre-ribosome before Ytm1, which was then incorporated in order to stabilize the binding (Rohrmoser et al., 2007).

An elegant study done in mammals showed that the amount of each protein in the cell was highly dependent on the other two components of the complex. Rohrmoser et al. demonstrated that any alteration in the amounts of free Bop1 (Erb1 in yeast) would exert an effect on the stability of its binding partners. It has been proposed that Bop1 interacts with Pes1 in order to shift to the nucleus and that Bop1 alone or Bop1-Wdr12 is trapped in the cytoplasm. Moreover, overexpression of Bop1 had a negative effect on cell growth probably because the excess of Bop1 in the cytosol would bind Wdr12 and prevent its transport to the nucleoplasm resulting in lack of PeBoW recruitment onto nascent ribosomes. This deleterious phenotype was rescued upon Wdr12 overexpression indicating the importance of the relative amounts of each protein for the proper function of the PeBoW. Interestingly, endogenous Bop1 would undergo a rapid degradation upon exogenous Bop1 overexpression. Therefore, Bop1 was proposed to play a central role in controlling the stability and degradation of the trimer, a fact that is line with the presence of PEST sequence in Erb1/Bop1 (Rohrmoser et al., 2007).

At last, several studies tried to explain which regions of the proteins were responsible for their physical interactions (Fig. 1.9). Final conclusions were made using different methods that allowed the co-purification and co-precipitation of the Nop7 sub-complex from yeast or PeBoW from mammals (Hölzel et al., 2007; Miles et al., 2005; Tang et al., 2008). Overexpression of mutants that lacked the N-terminal domain of Pes1 disrupted its incorporation into PeBoW and clearly showed

that this well conserved segment of the protein was required for the interaction. Ytm1-1 mutant in yeast and a set of mutants of Wdr12 that lacked WD40 repeats were shown not to bind Erb1 (or the association was severely destabilized), thus the β-propeller domain of Ytm1 was described to be involved in the interaction. It has been further confirmed by pull-down assays because GST-tagged N-terminally truncated Ytm1 was still able to co-elute with Erb1. In case of Bop1/Erb1 it was proposed that the central segment of the protein, close to the PEST sequence was sufficient for its interaction with Nop7 and Ytm1 but no direct evidence for the binding of this fragment of Erb1 to Nop7 and/or to Ytm1has been demonstrated (Tang et al., 2008). Moreover, it was suggested that the β-propeller domain of Erb1 was not involved in the Nop7 sub-complex assembly because initially it was described to be dispensable for Erb1 function.

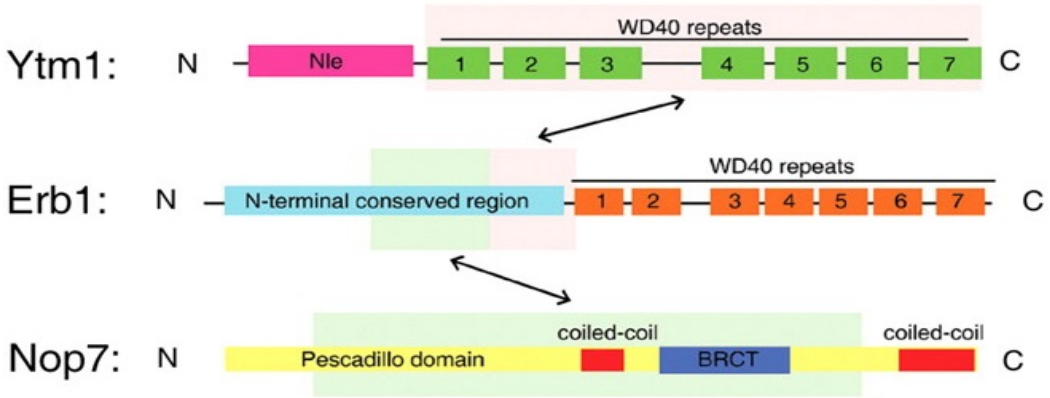


Figure 1.9 Summary of interactions proposed to occur between Nop7, Erb1 and Ytm1.
Erb1 is the core of the trimer and interacts with Nop7 through a region within its N-terminal domain. The exact site within Nop7 involved in binding to Erb1 is not yet known. The interaction of Ytm1 and Erb1 has been proposed to occur through the β-propeller domain of Ytm1 and a central segment of Erb1. Domain organization of each protein is shown. Green block depicts the segments responsible for Erb1/Nop7 interaction. Regions essential for Ytm1-Erb1 binding are shown with pink rectangles. (From Tang *et al.*, 2008).

Although the role of this large domain (44 kDa) of Erb1 in the context of ribosome biogenesis has remained unknown it is remarkable that two proteins within PeBoW complex carry a seven-bladed β-propeller. Moreover, other 20 AFs involved

in ribosome biogenesis contain WD40 repeats that fold into β-propeller structure and may mediate high-affinity protein-protein interactions that are extremely important in such a complex pathway (Henras et al., 2008).

1.3. WD40 domain

There are known many different structural motifs that appear in proteins and allow them to interact with each other. Among them, β-propeller domain formed by a repetition of segments called WD40 repeats is one of the most common (Stirnimann et al., 2010). Although it is not frequent in Bacteria, it became one of the most abundant protein domains in *S. cerevisiae* and remained widespread through the evolution (Chaudhuri et al., 2008). WD40 repeats are 40-50 residues long and often contain a conserved tryptophan-aspartic acid dipeptide (WD). Each repeat folds into 4 anti-parallel β-strands that eventually form blades of the propeller (Fig. 4.10) (Chen et al., 2011; Smith et al., 1999).

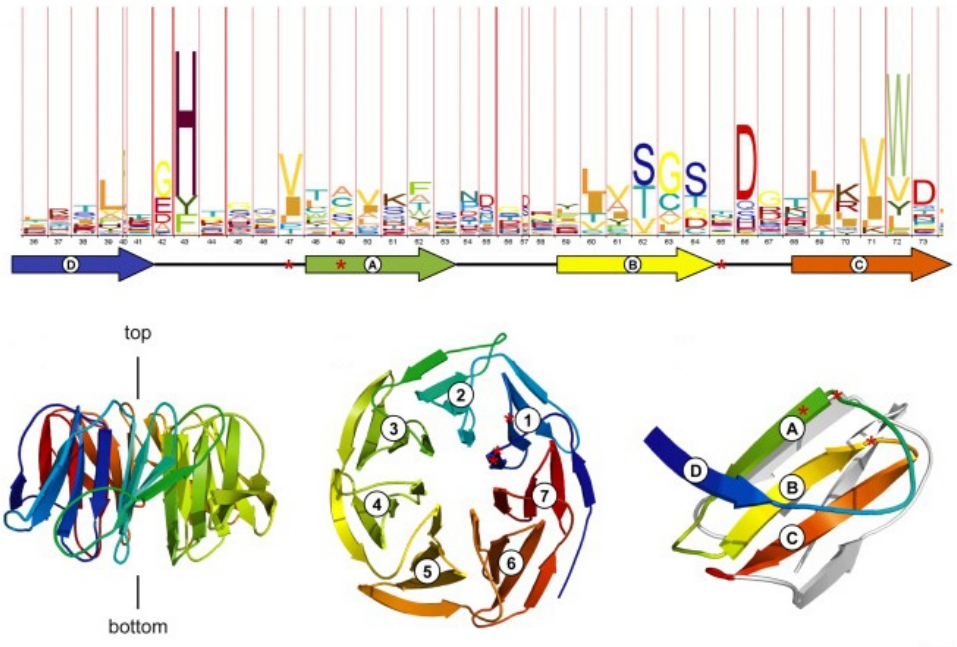


Figure 1.10 General features of WD40 repeats and β-propellers.
TOP: HMM logo of WD40 repeat representing the frequency of amino acids at given position within the motif and the corresponding β-strands that are formed upon folding. BOTTOM: An example of a seven-bladed β-propeller. The side view indicates the top/bottom orientation of the domain. Blade numbering is shown in the middle (top view). A canonical fold of a whole WD40 repeat demonstrates that strand D belongs to one blade whereas strands A, B and C form part of the following blade. (From Stirnimann et al., 2010).

By convention the first strand, called d, forms part of the blade n-1 whereas the next 3 strands a, b and c belong to the subsequent blade (n). The strands are connected by loops that can be very heterogeneous and, if long enough, may fold into additional motifs that flank the blades and contribute to the enormous diversity of the β-propellers (Xu and Min, 2011). The blades tend to close around an axis that remains hollow and becomes the central channel of the domain. Geometrical organization of the propeller allows to differentiate three well defined areas: the top face formed by loops connecting strands "d-a" and "b-c", the bottom face that includes loops "a-b" and "c-d", being both faces perpendicular to the central axis, and the circumference that comprises the lateral surface of the domain. From a functional point of view, the top face has been described as the one that most frequently interacts with binding partners, hence it contains a set of, so called, "hot spot" residues that are most likely to be involved in intermolecular interactions (Wu et al., 2012).

The number of WD40 repeats and, consequently, of the blades in a β-propeller is variable although members of WD40 family usually contain six, seven or eight blades. It is worth noting that there are other conserved families that fold into propeller-like structures but do not possess WD40 repeats. Historically it has been quite challenging to identify WD40 patterns in uncharacterized proteins because, in spite of a highly similar structure, the sequence of the domain is sometimes poorly conserved and in many cases the eponymous WD dipeptide is missing (Wang et al., 2013). Actually the most conserved feature of the family, at sequence level, is another aspartic acid residue (Garcia-Higuera et al., 1998). It appears in the "b-c" loops and is involved in maintaining blade organization by a network of electrostatic interactions with other conserved residues. It forms part of a structural triad that also involves a histidine that usually appears within a conserved GH motif in the loop "d-a" and a serine/threonine from strand "b". Triad formation is not indispensable for the correct folding of β-propellers but it is present in the vast majority of the cases.

Another particular feature of the domain is the way the circular structure is closed, in other words, how the N-terminal and C-terminal ends are brought together. The most common type of closure is achieved in a "velcro-like" manner

which leads to a formation of the last blade by strands coming from the first and the last WD-repeat. Often, the strand "d" of the first repeat joins strands "a", "b" and "c" from the last repeat and generates the closing blade (3+1 velcro). In several cases a 2+2 velcro or a closing that involves disulfide bridge formation have been described.

In general, the large surface of the domain, its compact structure and rigidity make it a perfect platform to establish high-affinity interactions with other macromolecules. Recently, it has been proposed that additional insertions or extensions that appear between the blades can modulate the binding to different ligands and confer the specificity of interactions (Xu and Min, 2011). The grade of conservation of the WD40 domains and their involvement in a wide range of molecular processes demonstrate the relevance and the versatility of β-propellers.

2. OBJECTIVES

Nop7-Erb1-Ytm1 trimer has been studied for the last 15 years and its relevance in eukaryotic ribosome biogenesis has been demonstrated. Striking phenotypes caused by depletion of the corresponding genes led to functional studies that have tried to elucidate the role of the complex in 60S assembly. However, only a limited amount of information has been obtained so far, probably due to an incredible level of complexity found in the pre-ribosomal complexes. In recent years a novel approach, based on the cryo-electron microscopy (cryo-EM), has been developed in order to study synthesis of the ribosome. Nevertheless, it requires high-resolution structures of the involved proteins that can be placed into cryo-EM models and then visualized in the context of the whole machinery.

Moreover, recent reports suggest that ribosome biogenesis plays crucial role in cancer cells because it provides means for high levels of protein synthesis that are required by fast proliferating cells. In fact, the genes coding for proteins of PeBoW complex (mammalian ortholog of Nop7-Erb1-Ytm1) over-express in certain types of tumors. Thus, one way of interfering with the development and growth of quickly dividing cells could be achieved by altering ribosome assembly pathway. In order to do so, a detailed knowledge about macromolecular interactions that occur during the process is fundamental.

Given the available partial information on the assembly of the complex and practically inexistent structural data regarding its architecture, the objectives of this work were:

1. Reconstitute Nop7-Erb1-Ytm1 trimer in vitro using heterologous expression systems
2. Study the properties of the individual proteins and the interactions that occur between them
3. Confirm previous findings regarding the segments of each protein involved in complex formation
4. Solve the structure of the components of the complex, their domains and the full trimer by means of X-ray crystallography
5. Corroborate any new structural information in the context of the cell
6. Investigate possible implications of altered association between Nop7, Erb1 and Ytm1 on cell proliferation and ribosome biogenesis.

3. MATERIAL AND METHODS

3.1. Cloning

Plasmids used in this study were generated using commercially available empty vectors: pOPIN vector suit was a gift from Dr. R. Owen's (OPPF, Oxford) and Dr. V. Rubio's (IBV, Valencia) labs, pET28-NKI/LIC 6His/3C vector was obtained from Dr. A. Perrakis group (NKI, Amsterdam), pNIC28-Bsa4 was received from Dr. A. Marina's lab (IBV, Valencia) and YCplac33/YCplac111 vectors were provided by Dr. Jesús de la Cruz (IBiS, Sevilla).

3.1.1. Preparation of pGKI vector

In order to switch from 6xHis-tag to GST-tag pET28-NKI/LIC 6His/3C vector was modified. NheI restriction site was generated by mutagenesis (detailed description of the procedure is available in Section 3.3), after the sequence coding for 6xHis-tag, using pNKI NheI F and pNKI NheI R oligos (Table 3.1). Resulting plasmid was digested with NcoI and NheI restriction enzymes (NEB) for 1h at 37ºC to remove the fragment corresponding to 6xHis-tag. GST-tag was amplified from pGEX-6P-2 vector using GST NcoI F and GST NheI R primers (Table 3.1). Resulting PCR product was then cloned into NcoI/NheI site of pET28-NKI/LIC backbone. Created vector, called pGKI allows the expression of proteins with N-terminal GST-tag that can be removed by 3C protease.

3.1.2. Preparation of constructs

The genes codifying proteins of interest were amplified by PCR using oligonucleotides listed in Table 3.1. The primers included overhangs specific for the chosen method of cloning (lower case letters in the sequences from Table 3.1) as described later. The PCR reaction volume was 25 µl and contained: 5 µl Reaction Buffer, 0.75 µl dNTPs mix (10mM each), Forward Primer 0.75 µl, Reverse Primer 0.75 µl, 1.25 µl DMSO, 0.25 µl of template DNA, 0.5 µl Kappa HiFi polymerase and 15.75 µl of MQ-water. PCR reaction was performed with the following setup:

3 min at 95ºC followed by 30 cycles of amplification (20 sec at 98ºC; 20 sec at desired Tm; 30 sec/kb at 72ºC) and finally the reaction was incubated for additional 5 min at 72ºC. Tm temperature and extension time varied depending on the oligonucleotides used and on the size of the amplicon.

Table 3.1 Oligonucleotides used to prepare the constructs described in this work.

The sequences of the overhangs compatible with corresponding vectors are shown in lower case.

Primer name	Sequence
NOP7 LIC F	5'-tacttccaatccatgAGAATCAAGAAGAAAAAC-3'
NOP7 LIC R	5'-tatccacctttactgCTATTTCTTGGAATCTAGTTT-3'
NOP7 NKI F	5'-cagggacccggtATGAGAATCAAGAAGAAAAACACCAGAG-3'
NOP7-465 NKI R	5'-cgaggagaagcccggttaACCAATAGCATCACCCCATGG-3'
NOP7-278 NKI F	5'-cagggacccggtACTGAGATTGAGGAAGACGTAAAAG-3'
ERB1 NKI F	5'-cagggacccggtATGATGGCTAAGAACAACAAAACTACCGAGG-3'
ERB1 NKI R	5'-cgaggagaagcccggttaGGTGGTCCATAAGCGAGCCGTATTATC-3'
YTM1 pOPINE F	5'-aggagatataccatgACAGAAGATAAATCGCAGGTTAAAATCAGG-3'
YTM1 pOPINE R	5'-gtgatggtgatgtttGTTTTTGAAAATGTTGTCTCCTTTATTTATTTGAATC-3'
YTM1 pOPIN F	5'-aagttctgtttcagggcccgACAGAAGATAAATCGCAGGTTAAAATCAGG-3'
YTM1 pOPIN R	5'-atggtctagaaagctttaGTTTTTGAAAATGTTGTCTCCTTTATTTATTTGAATC-3'
Cht NOP7NKI F	5'-cagggacccggtATGGGCAAGGCCAAGAAGAAGGG-3'
Cht NOP7 NKI R	5'-cgaggagaagcccggttaAGCCTTCTTCGCTGCCATTTCCTTC-3'
ChNOP7-350 NKI R	5'-cgaggagaagcccggttaGCCGCCGGGCGCG-3'
ChNOP7-350 NKI F	5'-cagggacccggtGGCGACGTACTCCCTCAACCC-3'
ChNOP7-466 NKI R	5'-cgaggagaagcccggttaGTTGATGCTGTCCCACACCCAC-3'
ChERB1 NKI F	5'-cagggacccggtATGGGGTCAAAAATAGTTGAAAAGAAGCG-3'
ChERB1 NKI R	5'-cgaggagaagcccggttaCATCCATAACCTCGCCGTCCCATC-3'
ChERB1-433 NKI F	5'-cagggacccggtCCCTCACCCGATGAGCTG-3'
ChYTM1 pOPIN F	5'-aagttctgtttcagggcccgATGGACGCCCCCATGGAGGACG-3'
ChYTM1 pOPIN R	5'-atggtctagaaagctttaTTTTTGCTCGGTAACAATATTTCTTC-3'
Erb1wt PstI F	5'-ccaagcttgcatgcCTGCAGGGCGAAATTTTTCTC-3'
Erb1wt BamHI R	5'-cggtacccggGGATCCAACCGCAATTACAGC-3'

PCR products were visualized on 1% agarose gels and purified using E.Z.N.A. Cycle Pure Kit (Omega bio-tek). Resulting DNA fragments were then ligated into corresponding plasmids using one of the three methods (as explained in Table 3.2)

- **Ligase Independent Cloning (LIC)**: Vectors were linearized by corresponding digestion with restriction enzymes and then both, vectors and inserts were treated separately with T4 DNA polymerase (NEB) in 20µl reaction mix that contained:

Vector	
T4 DNA polymerase Buffer	4µl
dTTP	0.5µl
DNA	15.3µl
T4 DNA Polymerase	0.2µl

PCR product	
T4 DNA polymerase Buffer	4µl
dATP	0.5µl
DNA	15.3µl
T4 DNA Polymerase	0.2µl

The mixtures were incubated 5 minutes at room temperature and then 10 minutes at 70ºC in order to inactivate the enzyme. Next, 2µl of vector and 2µl of insert were mixed and left 5 minutes at room temperature. Subsequently 1µl of 50mM EDTA was added and the reaction was transformed into chemically competent E. coli DH5α cells

- **In-Fusion Cloning** was performed using In-Fusion Kit from Clontech according to the manufacturer's manual and the resulting products were transformed into *E. coli* DH5α.
- **Sequence and Ligase Independet Cloning (SLIC)**: 2µl of vector linearized by restriction enzymes were mixed with 2µl of purified PCR product. Then 1µl of T4 DNA polymerase Buffer and 0.2µl of T4 DNA polymerase were added. The mixture was incubated for exactly 1 minute at room temperature and the immediately placed on ice and transformed into *E. coli* DH5α.

Table 3.2 summarizes the details of each construction described in this work and the ligation method used in each case.

The transformation in each case was carried out according to the following protocol:

5µl of the reaction mixture were added to 100µl of cells and left on ice for 20 minutes. The cells were shifted to 42ºC for 1.5 minutes and then placed on ice for another 2 minutes. 0.5ml of fresh LB was added at the end and the resulting mix was left for 1h at 37ºC on rocking platform. Cells were pelleted down and resuspended in 50µl of fresh LB and plated on LB-agar containing antibiotics compatible with the resistance provided by the plasmid. Successful cloning was confirmed by colony PCR and two positive colonies from each transformation were picked and used to inoculate 3ml of LB containing appropriate antibiotics which then was left over-night at 37ºC. The plasmid extraction was done using QuickGene-Mini80 device and

QuickGene Plasmid kit S from FujiFilm. Resulting constructions were sequenced and used for protein expression.

Table 3.2 Constructs prepared in the course of this work.

Primers as shown in Table 3.1 Method: Ligation Method used in each case.

Construct name	Primers used	Template	Vector	Method
Nop7-pNIC	NOP7 LIC F NOP7 LIC R	genomic DNA *S.cerevisiae*	pNIC28-Bsa4	LIC
Nop7_465-pNKI	NOP7 NKI F NOP7-465 NKI R	Nop7-pNIC	pET28-NKI/LIC 6His/3C	LIC
Nop7_278465-pNKI	NOP7-278 NKI F NOP7-465 NKI R	Nop7-pNIC	pET28-NKI/LIC 6His/3C	LIC
Erb1-pNKI	ERB1 NKI F ERB1 NKI R	genomic DNA *S.cerevisiae*	pET28-NKI/LIC 6His/3C	LIC
Ytm1-pOPINE	YTM1 pOPINE F YTM1 pOPINE R	Ytm1-pNKI	pOPINE	In-Fusion
Ytm1-pOPINF	YTM1 pOPIN F YTM1 pOPIN R	Ytm1-pNKI	pOPINF	In-Fusion
Ytm1-pOPINM	YTM1 pOPIN F YTM1 pOPIN R	Ytm1-pNKI	pOPINM	In-Fusion
Ytm1-pOPINS3C	YTM1 pOPIN F YTM1 pOPIN R	Ytm1-pNKI	pOPINS3C	In-Fusion
ChNop7-pNKI	Cht NOP7 NKI F Cht NOP7NKI R	genomic DNA *C. thermophilum*	pET28-NKI/LIC 6His/3C	LIC
ChNop7_350-pNKI	Cht NOP7NKI F ChNOP7-350 NKI R	ChNop7-pNKI	pET28-NKI/LIC 6His/3C	LIC
ChNop7_350466-pNKI	ChNOP7-350 NKI F ChNOP7-466 NKI R	cDNA *C. thermophilum*	pET28-NKI/LIC 6His/3C	LIC
ChErb1-pNKI	ChERB1 NKI F ChERB1 NKI R	cDNA *C. thermophilum*	pET28-NKI/LIC 6His/3C	LIC
ChErb1_432Ct-pNKI	ChERB1-433 NKI F ChERB1 NKI R	ChErb1-pNKI	pET28-NKI/LIC 6His/3C	LIC
ChYtm1-pOPINF	ChYTM1 pOPIN F ChYTM1 pOPIN R	cDNA *C. thermophilum*	pOPINF	SLIC
ChYtm1-pOPINJ	ChYTM1 pOPIN F ChYTM1 pOPIN R	cDNA *C. thermophilum*	pOPINJ	SLIC
Erb1-YCplac33	Erb1wt PstI F Erb1wt BamHI R	genomic DNA *S.cerevisiae*	YCplac33	SLIC
Erb1-YCplac111	Erb1wt PstI F Erb1wt BamHI R	genomic DNA *S.cerevisiae*	YCplac111	SLIC

3.2. Generation of cDNA library from *Chaetomium thermophilum*

Since several genes of interest from *Ch. thermophilum* contained introns it was necessary to prepare a library of cDNA in order to amplify those genes. The strain CDS144.50 was purchased from Fungal Biodiversity Center/CBS, Utrecht, The Netherlands. The fungus was grown in Eggins&Pugh medium as described in (Mushtaq and Jamill, 2012). Resulting mycelium was collected by centrifugation at 4000 x g at 4ºC.

First, total RNA was extracted from the mycelium using "SV total RNA Isolation Kit" from Promega with the following adaptations:

1. 100 mg ground mycelium were resuspended in 600 µl lysis buffer by pipetting 100 times up and down in a 2 ml test-tube
2. 1200 µl of dilution buffer were added and incubated for 3 minutes at 70°C in a water bath
3. The lysate was centrifuged 10 min / 20000 x g/ RT
4. The supernatant was split into two Eppendorf tubes (900 µl each)
5. 375 µl of 100% ethanol were added and mixed by inverting
6. Both solutions were transfered onto one Kit column (step by step)
7. 1000 µl of RNA wash solution were passed through the column
8. DNAse treatment was done as described in the manual
9. 75 µl nuclease free water were used to elute the RNA and it was stored at -80°C (measured concentration with Nano Drop was 1 µg/µl approximately)

The quality of the extracted RNA was assessed by agarose gel electrophoresis and resulting nucleic acid was used as a template for reverse transcription: Total reaction volume of 20 µl (2 reactions were carried out at the same time in order to produce higher amount of cDNA) contained: 7 µl total RNA (approx. 7µg), 1 µl Oligo d(T) primer (50mM), 4 µl dNTPs (2.5 mM) and 2 µl MQ-water. The reaction was incubated for 5 minutes at 65°C and then placed on ice for 1 minute. Then the mixture was completed with 4µl 5x Superscript Buffer (Invitrogen), 1µl 0.1 M DTT and 1µl Superscript-III-Reverse Transcriptase (Invitrogen)

Reverse transcription was done in a thermal cycler with the following setup: it was left for 5 minutes at 30°C followed by 5 hours at 52°C and finally it was

incubated for 15 minutes at 70ºC. When completed, the reaction was purified with E.Z.N.A. Cycle Pure Kit (Omega Bio-tek) and eluted cDNA was stored in Mili-Q water at -20ºC.

3.3. Site-directed mutagenesis

All mutants described in the present work have been generated by site-directed mutagenesis using oligonucleotides resumed in Table 3.3. Whole plasmid amplification by PCR was done in the following manner. Reaction mixture was set up as for insert amplification described in 3.1.2. Reaction conditions were: 3 min at 95ºC followed by 16 cycles of amplification (20 sec at 98ºC; 20 sec at desired Tm; 30 sec/kb at 72ºC) and finally the reaction was incubated for additional 5 min at 72ºC. The PCR products were visualized on 0.8% agarose gels and the template plasmid was removed by addition of 1.5 µl of DpnI and subsequent incubation at 37ºC for 1 h. 5 µl of the resulting mixture were used to transformed chemically competent *E. coli* DH5α. The transformation and plasmid extraction were carried out as described in 3.1.2. DNA was sequenced using suitable primers in order to confirm successful mutagenesis.

Table 3.3 Primers used in mutagenesis.
The sequence corresponding to the mutated codon is underlined.

Primer name	Sequence
ChERB E481R F	5'-GTCAGAGTGTGGAGACTGCTTACAGG-3'
ChERB E481R R	5'-CCTGTAAGCAGTCTCCACACTCTGAC-3'
ChE E481D F	5'-GTCAGAGTGTGGGATCTGCTTACAGGT-3'
ChE E481D R	5'-ACCTGTAAGCAGATCCCACACTCTGAC-3'
ChERB T484E F	5'-TGGGAACTGCTTGAAGGTCGGCAGG-3'
ChERB T484E R	5'-CCTGCCGACCTTCAAGCAGTTCCCA-3'
ChE T484Q F	5'-TGGGAACTGCTTCAAGGTCGGCAGGTT-3'
ChE T484Q R	5'-AACCTGCCGACCTTGAAGCAGTTCCCA-3'
ChERB R486A F	5'-GCTTACAGGTGCGCAGGTTTGGTC-3'
ChERB R486A R	5'-GACCAAACCTGCGCACCTGTAAGC-3'
ChE R486E F	5'-GCTTACAGGTGAGCAGGTTTGGTC-3'
ChE R486E R	5'-GACCAAACCTGCTCACCTGTAAGC-3'
ChE V488W F	5'-ACAGGTCGGCAGTGGTGGTCTGTCAA-3'
ChE V488W R	5'-TTGACAGACCACCACTGCCGACCTGT-3'

3.4. Protein expression in *Escherichia coli*

E. coli (DE3) BL21 CodonPlus (RIPL) was used to express all the constructions of Nop7, Erb1, ChNop7 and ChErb1. The plasmids of interest were transformed into chemically competent cells in a following manner:

0.5 µl of DNA was added to 50 µl of competent cells which were then incubated on ice for 20 minutes. A heat-shock pulse of 45 seconds at 42ºC was performed and the cells were placed back on ice for another 2 minutes. 1 ml of LB was added to the cells that were then incubated for 1 hour at 37ºC. 100 µl of the cells were plated on LB-agar containing appropriate antibiotics. The plates were left over-night at 37ºC.

The general procedure that was carried out in order to check soluble protein expression consisted in small-scale expression tests and small-scale affinity purification. If satisfactory amount of protein was recovered, expression and purification were scaled up.

3.4.1. Small-scale expression assays

Two colonies from *E. coli* (DE3) BL21 CodonPlus transformants were picked, one to be tested for expression in LB medium and the other one in ZY medium.

- IPTG expression test: 10ml of LB with antibiotics were inoculated with one colony from LB-agar plate and left for 8h at 37ºC. Then OD was measured and adjusted to 0.5 by addition of fresh LB medium. After another 30 minutes at 37ºC the expression of the protein of interest was induced with IPTG (final concentration of 0.5mM). The culture was left at 20ºC over-night.
- Auto-induction method: 10 ml of ZY containing $MgCl_2$, NPS, 5052 and antibiotics were inoculated with one colony of *E. coli* and were left at 37ºC until OD of 1 was reached. Then it was changed to 20ºC and left over-night.

In both cases, the cells were collected by centrifugation 4ºC / 4000 x g / 15 minutes and washed with 1ml of PBS.

In order to assess expression levels and solubility of the proteins, small-scale batch purifications were performed for each construction using Ni-NTA (for His-

tagged proteins), Glutathione Sepharose 4B (for GST-tagged proteins) or amylose (for MBP-tagged proteins) resins.

1. The pellet was resuspended in 300µL of the Lysis Buffer.
2. The sample was sonicated in a 1,5mL Eppendorf tube using UCD-200 Bioruptor (Diagenode) at medium intensity (200W) with 30 seconds ON/30 seconds OFF cycles during 15 minutes
3. The sample was then centrifuged for 4ºC / 21000 x g / 20 minutes in order to separate the soluble protein fraction
4. The supernatant (approx. volume 300µL) was then added to 50µL of beads which had been previously washed with 1mL of water and equilibrated with 1mL of corresponding wash buffer. (Ni-NTA beads: Buffer A (HS); Glutathione Sepharose 4B beads and amylose beads: Buffer SE)
5. The sample was left on a nutator at 4 ºC for 15 minutes to allow protein binding to the beads
6. The beads were then washed three times with 1mL of wash buffer each
7. Finally, 50µL of elution buffer were added in order to elute the protein

30µL of the sample were taken at each step and the presence of soluble protein was analyzed on by SDS-PAGE.

3.4.2. Large-scale expression

Those proteins that were successfully expressed at small-scale were subsequently scaled-up in order to carry out large-scale purification. Depending on the result of the beads-based purification either expression in LB induced with IPTG or ZY/auto-induction was chosen (Table 3.4 summarizes expression method used in each case).

- IPTG expression: 50ml of pre-culture were inoculated with corresponding strain and were left over-night at 37ºC. Next day, 950ml of fresh LB were mixed with the pre-culture and left at 37ºC until the OD was 0.6-0.8. At this point, 0.5ml of 1M IPTG were

added to the culture and the flasks were left at 20ºC over-night.

- Auto-induction: The large-scale protocol was adapted from Studier (Studier, 2005). Single colony was inoculated in 5ml of pre-culture containing ZY medium, 1mM Mg_2SO_4, 0.8% of glucose, 1x NPS and atibiotics. It was left over-night at 37ºC and subsequently it was transferred into 500ml of inducing media that contained: ZY, 1mM Mg_2SO_4, 1x 5052, 1x NPS and antibiotics. It was incubated at 37ºC and once it reached OD of 1 it was shifted to 20ºC and left over-night.

In each case, resulting cultures were centrifuged 4ºC / 3500 x g / 1h and the pellets were washed with 50ml of PBS, centrifuged 4ºC / 4000 x g / 30 minutes and then frozen in liquid nitrogen and stored at -80ºC until use.

Table 3.4 Summary of expression conditions and purification methods of the proteins used in this study.

Host column shows the organism that has been used for protein production (BL21 C+: *E. coli* (DE3) Codon Plus strain; Sf9: *Spodoptera frugiperda* cells). The last column indicates the purification steps that were performed for each protein and are explained in section 3.6 (His: Affinity chromatography of His-tagged proteins; Hep: Affinity chromatography using HiTrap Heparin HD column; GST: Affinity chromatography of GST-tagged proteins; SE: Size exclusion chromatography).

Protein	Residues	Tag (cleavage site)	Host	Expression method	Purification
Nop7	1-605	6xHis (TEV) – Nterm	BL21C+	0.5mM IPTG /20ºC /16h	His-SE
$Nop7_{1-465}$	1-465	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-SE
$Nop7_{278-465}$	278-605	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-SE
Erb1	1-807	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-SE-Hep
Ytm1	1-460	6xHis (3C) – Nterm	Sf9	Baculovirus / 72h	His-SE
Ytm1	1-460	6xHis (NO) – Cterm	Sf9	Baculovirus / 72h	His-SE
ChNop7	1-679	6xHis (3C) – Nterm	BL21C+	0.5mM IPTG / 20ºC/16h	His-SE
$ChNop7_{1-350}$	1-350	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-SE-Hep
$ChNop7_{350-466}$	350-466	GST (3C) – Nterm	BL21C+	Autoinduction	GST-SE
ChErb1	1-801	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-Hep-SE
$ChErb1_{432-801}$	432-801	6xHis (3C) – Nterm	BL21C+	Autoinduction	His-Hep-SE
ChYtm1	1-495	6xHis (3C) – Nterm	Sf9	Baculovirus / 72h	His-SE
ChYtm1	1-495	6xHis-GST (3C) - Nterm	Sf9	Baculovirus / 72h	GST-SE

3.5. Protein expression in insect cells using baculovirus.

Both, Ytm1 from *Saccharomyces cerevisiae* (ScYtm1) and Ytm1 from *Chaetomium thermophilum* (ChYtm1) could not be expressed in a soluble form in *E. coli*. It required a eukaryotic expression system and insect cells were the method of choice. A high-throughput method developed at OPPF (Oxford Protein Production Facility, UK) allowed to clone and test expression levels of ScYtm1 using pOPIN vectors suite. ScYtm1 was cloned into pOPINE (C-terminal, non-cleavable 6xHis-tag), pOPINF (N-terminal 6xHis-tag removable with 3C protease), pOPINM (N-terminal 6xHis-tag followed by an MBP-tag and 3C cleavage site) and pOPINS3C (N-terminal 6xHis-tag followed by SUMO-tag removable with SUMO-specific protease) plasmids and protein expression was tested in *Escherichia coli*, Sf9 cells and HEK293T cells (transient expression).

0.5µl of P_0 baculovirus stock of ScYtm1 (N- or C-terminally His-tagged) produced at OPPF (UK) were amplified in 30ml of Sf9 cells ($1.5x10^6$cells/ml) grown in Sf900II medium (Life Technologies) at 27ºC for 5 days. Resulting P_1 virus were used for protein expression at large-scale. 10ml of P_1 stock were added to 300ml of Sf9 cells at $1.5x10^6$cells/ml and the culture was left at 27ºC for 72h. Infected cells were collected by centrifugation at 4ºC / 1000 x g / 15 minutes and washed with PBS with a subsequent centrifugation step with the same setup. Resulting pellets were frozen in liquid nitrogen and kept at -80ºC until use.

ScYtm1-pOPINJ, ChYtm1-pOPINJ and ChYtm1-pOPINF were generated by In-Fusion cloning as previously described and used for Sf9 cells transfection as follows:

1. Ian Jones bacmid (Zhao et al., 2003) preparation:
    - The bacmid was extracted from 5ml *E. coli* culture using Qiaprep Spin Miniprep Kit (Qiagen)
    - 150 µg of bacmid were digested with Bsu36I for 3h at 37ºC and then incubated for 20 minutes at 72ºC to inactivate the enzyme
    - aliquots of 1 µg of bacmid were prepared and stored at -20ºC until use

2. 2 ml of Sf9 cells were disposed on a six-well plate ($1x10^6$ of cells/well) and the plate was left at 27ºC for 1h
3. Transfection mix was prepared (per well): 200 µl of Sf900II medium, 1 µg of Ian Jones bacmid, 1 – 1.5 µg of pOPIN-derived plasmid and 4 µl of FugeneHD (Promega) and was left for 30 minutes at room temperature
4. Transfection mix was added to each well containing Sf9 cells
5. The plate was left at 27ºC and was visually inspected each 24h.
6. Approximately 5-6 days after transfection, the supernatant containing P_0 virus was collected from each well

The cells from each well were lysed using 100 µl of FastBreak Cell Lysis Reagent (Promega) and the lysate was centrifuged at 4ºC / 21000 x g / 15 minutes in order to separate the soluble fraction. Both, soluble and insoluble fractions were then loaded on an SDS-PAGE gel and then blotted on a nitrocellulose membrane that was subsequently incubated with Anti-6xHis Antibody (Roche) in order to assess successful transfection by detection of recombinant proteins.

P_0 virus were further amplified and used for large-scale protein expression in 300ml of Sf900II medium as described previously.

3.6. Protein purification

The composition of the buffers used in different purification steps is listed in Table 3.5.

3.6.1. Protein extraction

Soluble proteins were recovered in a similar manner for both, E. coli and Sf9, cultures. Frozen cell pellet was thawed on ice and resuspended in 30ml of cold Lysis Buffer supplemented with one pill of Complete EDTA-free Protease Inhibitor Cocktail (Roche) and 10mM imidazole (His-tagged proteins only). The cells were lysed by sonication during 30 minutes (15 minutes in case of Sf9 cells) (1 second ON + 1 second OFF pulses) using W75042 Sonicator (Fisher Scientific). The sample

was centrifuged for 35 minutes at 17000rpm (Sorvall SS-34 Rotor) at 4ºC and resulting supernatant was passed through 0.45µm filter (VWR).

Table 3.5 Buffers used in protein purifications.
Low-salt buffers (LS) are identical as high-salt buffers (HS) except for NaCl concentration. LS buffers were used when Heparin HiTrap column was included in the purification scheme.

Buffer Name	Component		Buffer Name	Component	
Lysis Buffer	Hepes pH 7.5	50 mM	**Buffer A (LS)**	Hepes pH 7.5	50 mM
	NaCl	500 mM		NaCl	150 mM
	glycerol	10.00%		glycerol	5.00%
	Triton X-100	0.10%		imidazole	20 mM
	B-mercaptoethanol	5 mM		B-mercaptoethanol	2 mM
Buffer A (HS)	Hepes pH 7.5	50 mM	**Buffer B (LS)**	Hepes pH 7.5	50 mM
	NaCl	500 mM		NaCl	150 mM
	glycerol	5.00%		glycerol	5.00%
	imidazole	20 mM		imidazole	500 mM
	B-mercaptoethanol	2 mM		B-mercaptoethanol	2 mM
Buffer B (HS)	Hepes pH 7.5	50 mM			
	NaCl	500 mM			
	glycerol	5.00%			
	imidazole	500 mM			
	B-mercaptoethanol	2 mM			
Buffer SE	Hepes pH 7.5	20 mM			
	NaCl	150 mM			
	glycerol	5.00%			
	B-mercaptoethanol	2 mM			
Buffer BGH	Hepes pH 7.5	20 mM			
	NaCl	150 mM			
	glycerol	5.00%			
	reduced gluthathione	20 mM			
	B-mercaptoethanol	2 mM			

3.6.2. Affinity chromatography of His-tagged proteins.

IMAC chromatography was performed using 5ml HisTrap FF (GE Healthcare) columns. Initially, the column was washed with 10 colum volumes (CVs) of water and then it was equilibrated with 10 CVs of Buffer AHS. Filtered lysate was then loaded on the column. Next, additional 40 ml of Buffer AHS were passed in order to wash the column. The protein of interest was eluted from the column using FLPC system ÄKTAPurifier (GE Healthcare). Increasing

concentrations of Buffer BHS initially allowed to remove contaminants that attach to Ni2+ with low affinity. Eluted fractions were checked on SDS-PAGE and those that contained protein were mixed and concentrated using Amicon Ultra (Milipore) with 30 kDa or 50 kDa cutoff until the volume was of 1 ml.

Described procedure was slightly modified for proteins that tend to bind nucleic acids. After they were loaded on the column and washed with 40ml of Buffer AHS, a second step of wash was done using buffer with lower amount of NaCl (Buffer ALS). Elution was performed with buffer BLS so that the osmotic pressure would not interfere with the subsequent purification on HiTrap Heparin column. Eluted fractions that contained protein of interest were mixed.

3.6.3. Affinity chromatography of GST-tagged proteins.

Cleared lysate was loaded on a 5 ml GSTrap FF column (GE Healthcare) previously washed with 10 CVs of water and another 10CVs of Buffer SE. 30 ml of Buffer SE were passed through the column in order to remove unbound proteins. The elution was done using Buffer GSH. Eluted fractions were then analyzed on a 10% SDS-PAGE and those that contained GST-protein were concentrated in a 50 kDa Amicon Ultra. When the removal of GST tag was convenient the fractions containing GST-tagged protein were pooled, Prescission Protease (PP) was added to a final concentration of 1 μg/ml and the mixture was left over-night at 8ºC. Then it was passed again through GSTrap column in order to remove uncleaved protein, GST tag alone and PP protease. The flow through of the column was then concentrated and injected in gel filtration column.

3.6.4. Affinity chromatography using HiTrap Heparin HD column.

Proteins that co-purified with nucleic acids were subjected to additional purification steps after they were eluted from HisTrap. The column was washed with 10CVs of water and the equilibrated with 10CVs of Buffer SE. The sample containing protein eluted from HisTrap was loaded on the column which was ten washed with 30ml of Buffer SE. The elution was achieved increasing NaCl concentration (Buffer BHEP) in a linear manner. The absorbance at 260 and 280 nm was constantly monitored in order to detect nucleic acid-free protein fractions that

were then checked on a SDS-PAGE gel. Fractions containing the protein were mixed and concentrated up to 1 ml.

3.6.5. Size exclusion chromatography (gel filtration).

After affinity chromatography, concentrated protein was injected into Superdex 200 16/60 column (GE Healthcare) previously equilibrated with Buffer SE. Elution of the protein was monitored by absorbance at 280 nm. Fractions corresponding to the peaks were run on a 10% SDS-PAGE in order to confirm protein presence. Those containing pure monomeric protein were mixed and concentrated using Amicon Ultra until the volume was below 500 µl (depending on the amount and solubility of each protein). The concentration of the protein was measured using Bradford reagent (Bio-Rad). The sample was then aliquoted and flash-frozen in liquid nitrogen.

3.7. Protein stability assays

In order to check thermal properties of the proteins and to confirm enhanced stability of those corresponding to *Chaetomium thermophilum* Thermofluor assay (Pantoliano et al., 2001) was performed as follows: The protein was diluted to the final concentration of 1mg/mL in Buffer SE up to 198µL of the total sample volume. 2µL of SyproOrange Protein Gel Stain (Sigma) were added. 4µL of the mix were added to each of the 48 wells of the PCR plate (Rows A to D) (Table 4.6). 20µL of the sample from each well from rows A to D were transferred to the corresponding wells in rows E to H. The plate was sealed with a plastic film and it was centrifuged 1 minute at 4000rpm/4ºC. The plate was then analyzed in the Fast 7500 Real Time PCR system (Appiled Biosystems). The run was set up to start at 20ºC and the final temperature reached after 65 cycles was 85ºC; the fluorescence emission was measured at 550 nm. The data were analyzed using GraphPad Prism software and the wells with the highest V50 score were considered to contain the optimal conditions to enhance the protein stability.

Table 3.6 Schematic organization of the conditions tried in Thermofluor assays.
The variable component of each condition is shown in red

	1	2	3	4	5	6	7	8	9	10	11	12
A	150mM NaCl 100mM AcNa pH 4.5	150mM NaCl 100mM Bis Tris pH 5.5	150mM NaCl 100mM citrate pH 5.6	150mM NaCl 100mM cacodilate pH 6.5	150mM NaCl 100mM MES pH 6.5	150mM NaCl 100mM Hepes pH 7	150mM NaCl 100mM Hepes pH 7.5	150mM NaCl 100mM Fosfato pH 7.5	150mM NaCl 100mM Tris pH 8	150mM NaCl 100mM Tris pH 8.5	150mM NaCl 100mM Bicina pH 9	150mM NaCl 100mM CHES pH 9.5
B	150mM NaCl 100mM SPG 4.5	150mM NaCl 100mM SPG 5.5	150mM NaCl 100mM SPG 6.5	150mM NaCl 100mM SPG 7.5	150mM NaCl 100mM SPG 8	150mM NaCl 100mM SPG 9	150mM NaCl 100mM MMT4.5	150mM NaCl 100mM MMT5.5	150mM NaCl 100mM MMT6.5	150mM NaCl 100mM MMT7.5	150mM NaCl 100mM MMT 8	150mM NaCl 100mM MMT 9
C	Hepes pH 7.5	100mM NaCl Hepes pH 7.5	250mM NaCl Hepes pH 7.5	500mM NaCl Hepes pH 7.5	150mM NaCl 100mM Hepes pH7.5 5% Glicerol	150mM NaCl 100mM Hepes pH7.5 10% Glicerol	150mM NaCl 100mM Hepes pH 7.5 15% Glicerol	150mM NaCl 100mM Hepes pH 7.5 1mM MgCl2	150mM NaCl 100mM Hepes pH 7.5 1mM CaCl2	150mM NaCl 100mM Hepes pH 7.5 0.1mM ZnCl2 0.1mM CoCl2	150mM NaCl 100mM Hepes pH 7.5 1mM CdCl2 1mM MnCl2	150mM NaCl 100mM Hepes pH 7.5 1mM DTT
D	150mM NaCl 100mM Hepes pH7.5 1mM EDTA	100mM Hepes pH7.5 150mM LiCl2	100mM Hepes pH7.5 150mM KCl	100mM Hepes pH7.5 150mM NaF	100mM Hepes pH7.5 150mM SO4NH4	100mM Hepes pH7.5 150mM NO3NH4	150mM NaCl 100mM Hepes pH 7.5 100mM Sacarosa	150mM NaCl 100mM Hepes pH 7.5 300mM Sacarosa	150mM NaCl 100mM Hepes pH 7.5 Urea 0.5M	150mM NaCl 100mM Hepes pH 7.5 L-Arg 50mM L-Glu 50mM	150mM NaCl 100mM Hepes pH 7.5 Betaina 100mM	150mM NaCl 100mM Hepes pH 7.5 1%TritonX-100

3.8. Native gel electrophoresis

Total amount of 20 µg of protein was mixed with 5X Sample Buffer (0.3M Tris-HCl pH 8.5, 0.05% (w/v) bromophenol blue, 40% (v/v) glycerol) and loaded in an 8% native polyacrylamide gel. The gel vas previously pre-ran for 40 minutes in Tris-Glycine running buffer at pH 8.3. The samples were run on the gel at 4ºC for 3h and the voltage used was 100V.

3.9. Measurement of binding in vitro

3.9.1. Poly(U)-agarose beads binding.

Approximately 20µl of the polyuridylic acid-agarose (polyU) were used by assay. The beads were equilibrated 5 times with 500µl of reaction buffer each

(50mM Tris pH 8; 100mM NaCl; 5mM MgCl2; 3mM DTT; 0.1% Triton-X100 and 0.1 mg/ml BSA) then 500µl of the sample containing 200µg of protein diluted in the reaction buffer were loaded on the beads and incubated at room temperature for 30 minutes. Unbound protein was removed by washing the beads five times with wash buffer (50mM Tris pH 8; 100mM NaCl; 5mM MgCl2; 3mM DTT; 0.1% Triton-X100). The bound fraction was eluted by addition of 30µl of 6x loading buffer and boiling for 5 minutes at 95°C. The samples were then analyzed on 10% polyacrylamide gel. The saturation was studied by incubating the protein sample with 0.1mg/ml or 1mg/ml of free polyuridylic acid for 30 minutes before it was loaded onto the beads.

3.9.2. Electrophoretic mobility shift assay (EMSA)

30ng of dsDNA (120 base pairs) were incubated alone or with different amounts of $ChErb1_{432-801}$ in SE Buffer for 30 minutes on ice. Then the samples were mixed with 6X Sample Buffer and run on 1% agarose gel at 4ºC to avoid protein denaturalization. In order to visualize nucleic acids 1X GelRed (Biotium) was added to the gel.

3.9.3. Pull-down.

100µg of bait protein in Buffer SE were loaded onto 50µl of beads equilibrated with the same buffer. After 15 minutes of incubation on ice the beads were extensively washed with Buffer SE and equimolar amounts of putative binding partner were added to the beads and left on ice for 15 minutes. Free protein was removed in three wash steps (1ml of Buffer SE each) and the bound fraction was eluted with corresponding elution buffer. Samples from the input, final wash and elution steps were analyzed on SDS-PAGE gel. For 6xHis-tagged proteins NTA Agarose Beads (ABT) were used, Sepharose Glutathione 4B beads (GE Healthcare) were chosen for GST-tagged proteins and in case of MBP-tagged molecules Amylose Resin from NEB was applied.

3.9.4. Bio-layer Interferometry.

Biolayer interferometry system (BLItz, PALL) was used to calculate KD values for the interaction between ChYtm1 and ChErb1 or ChErb1432-801 versus

ChErb1 mutants. A sample containing 40ng/µl of GST-ChYtm1 in SE Buffer was loaded onto Anti-GST Biosensor (ForteBio) during 180s which was followed by 60s of equilibration in Buffer SE. Next, association and dissociation steps (180s each) were carried out using different concentrations of binding partners also diluted in SE Buffer. Curve fitting and KD calculation were done using BLItz Pro 1.2 Software.

BioLayer Interferometry also served to determine dissociation constant for the interaction between ssRNA and ssDNA and the β-propeller domain of ChErb1. A sample containing 50µg/ml of 15 nucleotide-long 5'-biotinylated polyU or 20 nucleotide-long 5'-biotinylated DNA from Sigma-Aldrich was immobilized on Streptavidin biosensors (Forte Bio) previously hydrated with sample buffer (50mM Hepes pH 7.5; 150mM NaCl; 5% glycerol and 2mM β-mercaptoethanol). Increasing amounts of $ChErb1_{432-801}$ (were used in association and dissociation steps. Curve fitting of triplicates and K_D calculation were carried out with BLItz Pro 1.2 software.

3.9.5. Isothermal Titration Calorimetry (ITC).

Nano ITC from TA Instruments was used to perform the measurements. The experiments were done at 20ºC and consisted of 24 serial injections of 1.5µl of 280µM 6xHis-ChYtm1 into the cell compartment containing 170µl of 23µM 6xHis-$ChErb1_{432-801}$. Both samples were diluted in Buffer SE. The heat was corrected by subtracting heat of dilution measured by injecting 280µM 6xHis-ChYtm1 into the cell compartment containing 170µl of SE Buffer only with the same experimental setup. Base-line correction was done with Nitpic and for data fitting and binding kinetics calculation Sedphat software was used (Houtman et al., 2007; Keller et al., 2012). The graphs were generated with GUSSI .

3.10. Protein crystallization

3.10.1. Crystals of $Erb1_{416-807}$

After size exclusion chromatography, concentrated samples of Nop7 and Erb1 were mixed in equimolar amounts and injected into Superdex200 16/60 column equilibrated with SE Buffer. The 5ml fractions corresponding to the heterodimer were mixed, concentrated and used for crystallization. Initial crystallization trials were performed at 21ºC, the concentration of Nop7-Erb1

complex was 80mg/ml and drops containing 0.3µl of protein sample and 0.3µl of reservoir were set up. Crystals diffracting up to 2.9Å were obtained in 0.1M Hepes pH 7.5; 10% Polyethylene glycol 8000 and 8% ethylene glycol. In order to improve crystal size and resolution we performed an optimization screen based on the Hampton Additives kit, the protein concentration used was 60mg/ml and the drop size of 0.5µl for protein sample, 0.5µl for reservoir and 0.1µl of additive was added. Crystals obtained in the original screening and from the Hampton Additives screening were flash-cooled in liquid nitrogen and diffracted at Diamond Light Source (Harwell, UK) I24 and I03 beamlines. The maximum diffraction up to 1.6Å was obtained with crystals grown with addition of 30% of ethanol. The crystals contained one molecule per asymmetric unit and the space group was P 21 21 21 with the following unit cell parameters: a = 52.026Å, b = 62.432Å, c = 158.22Å; $\alpha=\beta=\gamma = 90.00^{o}$. The data were processed using XDS (Kabsch, 2010), merged and scaled in CCP4 (Winn et al., 2011). Molecular replacement was done in parallel using MR module of Phenix (Adams et al., 2010) and Balbes on-line MR suite (Long et al., 2008). The search model used in phenix.mr was a poly-Ala β-propeller based on the input model chosen by Balbes database search (PDB: 2H13), which corresponds to an unrelated WD40 protein, WDR5, engaged in histone binding. Obtained model and initial phases were then used for additional model building cycle by AutoBuild module from Phenix. The final structure was then refined combining phenix.refine suite and manual refining in Coot (Emsley et al., 2010) until the final r factors were R = 16.0% and R free = 17.4%. Data collection and refinement statistics are shown in Table 1. The model and structure factors were deposited in Protein Data Bank with PDB ID: 4U7A.

3.10.2. Crystals of ChErb1$_{432-801}$-ChYtm1

ChErb1$_{427-801}$-ChYtm1 dimer in P6$_5$ 2 2 space group. Hexagonal crystals were obtained upon crystallization trials of ChErb1 full-length mixed with ChYtm1 (1:1 ratio, final concentration 38 mg/ml). Initial sitting-drop crystallization screenings were performed at 294K. The best resolution was achieved for crystals that grew in 15% PEG 4000; 0.1M sodium citrate pH 5.6; 0.2M Ammonium Sulfate that diffracted up to 3.2Å in XALOC beamline of ALBA Synchrotron (Spain). Data

were processed using iMosflm (Leslie and Powell, 2007) and CCP4 and the phasing was done using a combined Molecular Replacement approach. In the first step Balbes server was used to place the β-propeller of ChYtm1. Resulting model, altogether with the PDB of previously solved $Erb1_{416-807}$ from yeast, were used as input for Phaser-MR module from Phenix. Next, AutoBuild from Phenix partially placed residues corresponding to both propellers. The final model was completed by manual fitting of missing loops and regions of the UBL domain of ChYtm1. Cycles of phenix.refine and manual refinement were performed in order to obtain the final structure with its R and R_{free} factors of 19.4 and 23.9 (%) respectively

$ChErb1_{432-801}$-ChYtm1 dimer in $P2_1\ 2_1\ 2$ space group. Orthorhombic crystals appeared in sitting drops that contained ChNop7-ChErb1-ChYtm1 trimer (35 mg/ml) and the best diffracting crystals were obtained in 20% PEG 8000 and 0.1M Hepes pH 7.5 at 294K. The crystals were cryo-protected by adding 10% ethylene glycol to the crystallization condition and flash-cooled in liquid nitrogen. The data were collected at100K at I04 beamline at DLS facility (UK) where automated data processing was performed by xia2 pipeline (Winter, 2009). The highest resolution dataset (2.1Å) was used for phasing with Phaser-MR from Phenix and the refined 3.2Å structure served as the input model. The resulting model was refined and the position of UBL domain of ChYtm1 was manually corrected in Coot. The final structure was refined manually and with phenix.refine, until the values of R and R_{free} of 16.4% and 21.3% were reached.

$ChErb1_{432-801}$[R468]-ChYtm1 dimer in $P6_5\ 2\ 2$ space group. Crystals grew in a drops containing 0.3µl of ChErb1R486$E_{432-801}$/ChYtm1 at 6 mg/ml with 0.3µl of 0.1M Hepes pH 7.5 and 2M ammonium sulfate. They were flash-cooled in liquid nitrogen and diffracted at 100K at I02 Beamilne of DLS Synchrotron (UK). Acquired data were automatically processed and scaled with xia2. Phenix.refine was used for phasing and the $P6_5$ 2 2 structure of $ChErb1_{427-801}$/ChYtm1 was used as input. Resulting model was manually inspected with Coot and refined using phenix.refine. Final R and R_{free} values were 20.2% and 26.2%, respectively.

3.11. Bioinformatics tools used for structure and sequence analyses and representation.

PISA Server was used to assess biological assembly of the PDB models in solution and to compute buried area of the dimer (Krissinel and Henrick, 2007). CONTACT module from CCP4 served to identify residues involved in dimer formation. Similarity searches were done against PDB Database using Dali webserver and iPBA server (Gelly et al., 2011; Holm and Rosenström, 2010). Electrostatic potential was assigned by APBS tool in Chimera (Baker et al., 2001; Meng et al., 2006). Superimposition of structures and RMSD calculation was performed with MatchMaker module from Chimera. Chimera suite was also used to generate structure figures. Sequence multi-alignments were done using T-coffee webserver tool integrated in JalView (Notredame et al., 2000; Waterhouse et al., 2009), which was also used to depict protein sequences. WDSP server was used to predict WD40 repeats and hot-spot residues in Erb1 and Ytm1 (Wang et al., 2013). Docking of polyA into Erb1Ct structure was performed with HADDOCK server (de Vries et al., 2010).

3.12. Functional assays in *S. cerevisiae*

3.12.1. Yeast strains and plasmids used.

The strains discussed in this study have been prepared in the laboratory of Dr. Jesús de la Cruz at Instituto de Biomedicina de Sevilla (IBiS), Seville, Spain.

BY4741 was used as the parental strain (Brachmann et al., 1998). Strain JDY1232 (BY4741 but erb1::kanMX4) is a haploid segregant of Y26184 (Euroscarf) that requires a plasmid-borne copy of erb1 for cell viability.

To generate YCplac111-Erb1 and YCplac33-Erb1, a 4150 kb PstI/BamHI fragment containing the Erb1 ORF and 1000 bp up- and downstream the ATG and stop codons, respectively, was amplified from genomic DNA of the BY4741 strain and cloned into the PstI/BamHI-restricted vectors YCplac111 and YCplac33 (for details regarding oligonucleotides and cloning procedure check section 3.1.2). Both constructs apparently complemented the erb1 null allele to the wild-type extent.

The mutant erb1[R470E] was generated by site-directed mutagenesis in YCplac111-Erb1 with appropriate primers (Table 3.1) and confirmed by sequencing. This construct and YCplac111-Erb1 were transformed into the shuffle strain JDY1232-YCplac33-ERB1 and subsequently the YCplac33-Erb1 was counter-selected on plates containing 5-fluoroorotic acid (5-FOA).

3.12.2. Yeast transformation

For each transformation to be done 1 ml of corresponding medium was inoculated with yeast and was grown over-night until the OD was 0.9-1. The cells were pelleted, washed with sterile water and pelleted again. 100 µl/transformation of Solution I (10mM Tris-HCl pH 7.5, 1mM EDTA pH 8, 100mM lithium acetate) were added to the cells. For each transformation 5 µl of salmon sperm (10mg/ml) were dispensed in an Eppendorf tube and it was mixed with approximately 1 µg of plasmid. Then 100 µl of cells were added to each tube and the mixture was vortexed. Next 600 µl of Solution II (10mM Tris-HCl pH 7.5, 1mM EDTA pH 8, 100mM lithium acetate, 40% PEG 3350) were added and gently mixed with the cells. Transformation reaction was incubated for 30 minutes at 30ºC in a thermomixer and then it was kept at 42ºC for 15 minutes. Finally, the cells were pelleted down, resuspended in 80 µl of sterile water and plated on selective media.

3.12.3. Protein extraction from yeast

Extraction of proteins from yeast was done as follows: 20ml of suitable medium were inoculated with one colony and grown until log phase. The cells were centrifuged for 5 minutes at 4000 rpm, washed with water and then pelleted again. 100 µg of glass beads and 100 µl of Lysis Buffer (50mM Tris-HCl pH 7.5, NaCl 100mM, 1.5mM $MgCl_2$, 0.15% NP40) were added to the pellet. The cells were lysed during 40s using FastPrep. Resulting lysate was centrifuged for 10 minutes at 13000 rpm/4ºC in order to separate soluble proteins from the glass beads and cell debris. Protein concentration was measured in the supernatant which was then used for western blotting.

3.12.4. Immunoblotting

Purified Nop7, Erb1 and Ytm1 were used for rabbit immunization at "Centro de Producción y Experimentación Animal" facility (Sevilla, Spain) in order to produce antibodies. Gel electrophoresis of 10% SDS-PAGE was done during 1h at 100V, next the proteins were transferred to nitrocellulose membrane for 2h at 80V. The membranes were blocked with 5% (w/v) nonfat dry milk for 1h, incubated with the primary antibody (1:10 000 dilution) over-night, washed with TBST and incubated with Goat Anti-Rabbit antibody (GE Healthcare, 1:5000 dilution) for one hour. Subsequently the membranes were washed with TBST and the presence of HRP conjugated antibodies was detected using ECL Western Blotting Substrate (Pierce).

3.12.5. Polysome profiling.

200ml of YPD medium were inoculated with yeast and were grown until the OD was 0.5. 1ml of cyclohexamide was added to the culture and it was left on ice for 5 minutes. Next, the cells were collected by centrifugation for 5 minutes / 5 000 x g / 4ºC and washed with 20 ml of Lysis Buffer (10mM Tris-HCl pH 7.5, 100mM NaCl, 30mM $MgCl_2$, 100 µg/ml cyclohexamide, 200 µg/ml heparin, 0.2 µl/ml DEPC) and pelleted again for 5 minutes / 5 000 x g / 4ºC. Finally, the pellet was resuspended in 1ml of Lysis Buffer and transferred to a 1.5 ml Eppendorf tube and spun down for 1 minute. Approximately 400 µl of Lysis Buffer and the same volume of glass beads were added to the cells that were then disrupted by vortexing for 8 minutes at 4ºC. The lysate was separated from the beads by filtration and the insoluble debris were removed by centrifugation for 10 minutes / 13 000 x g /4ºC. The supernatant was collected and flash-frozen in liquid nitrogen upon use. Polysome extracts were loaded on 10-50% sucrose gradient and ultra-centrifuged for 2h 45' / 39 000 x g / 4ºC. Gradient analysis was performed using ISCO UA-6 gradient UV-detection and gradient collection system with continuous monitoring at A_{254}. If required 500 µl fractions were collected from the gradients and then subjected to trichloroacetic acid (TCA) protein precipitation. Briefly, 50 µl of TCA were added to each 500 µl fraction. After centrifuging for 10 min at 4ºC / 13 000 x g, the pellets were washed twice with cold acetone. Finally, Sample Buffer was

added to the pellet and it was boiled at 95ºC for its subsequent use in SDS-PAGE and Western Blotting.

The following methods were not directly performed by PhD candidate but in collaboration with Dr. Jesús de la Cruz and the results of the experiments are presented as a part of this work.

3.12.6. Ribosomal subunits fractionation

Cells were grown as in 3.12.5. They were treated with 1mM of NaN3 for 20 minutes and cyclohexamide was omitted. The lysis buffer used in this case was 50mM Tris pH 7.5, 50mM NaCl and 1mM DTT. Cell lysis and subsequent steps were performed as in 3.12.5 with the only difference being the ultra-centrifugation time which in this case was 4h.

3.12.7. RNA extraction, northern blot and primer extension.

RNA extraction, northern blot hybridization and primer extension analyses were carried out according to standard procedures (Venema et al., 1998). In all experiments, RNA was extracted from samples corresponding to 10 OD_{600} units of exponentially grown cells. Equal amounts of total RNA (5 μg) were loaded on gels or used for primer extension reactions. Specific probes, whose sequences are described in (Rosado and de la Cruz, 2004), were 5'-end labelled with [γ32-P] ATP. Signal intensities were quantified using a FLA-5100 imaging system and Image Gauge (Fujifilm).

The following oligonucleotides were used: 5′A_0 (5′-GGTCTCTCTGCTGCCGG-3′), D/A_2 (5′-GACTCTCCATCTCTTGTCTTCTTG-3′), A_2/A_3 (5′-TGTTACCTCTGGGCCC-3′), E/C_2 (5′-GGCCAGCAATTTCAAGTTA- 3′), C_1/C_2 (5′-GAACATTGTTCGCCTAGA-3′), 18S (5′-CATGGCTTAATCTTTGAGAC-3′), 5.8S (5′-TTTCGCTGCGTTCTTC ATC-3′), 5S (5′-GGTCACCCACTACACTACTCGG-3′)

4. RESULTS

4.1. Reconstitution of Nop7-Erb1-Ytm1 from *S. cerevisiae*

4.1.1. Expression and purification of Erb1

6xHis-Erb1 (94 kDa) was successfully expressed in Escherichia coli (DE3) BL21 CodonPlus strain using autoinduction protocol. Soluble protein was purified using HisTrap column followed by size exclusion chromatography and HiTrap Heparin (Fig. 4.1). The protein co-purified with large amount of nucleic acids that were removed in Heparin purification step. Large amounts of bacterial proteins were co-purifying with Erb1 after HisTrap column but they were efficiently removed by HiTrap Heparin. In gel filtration two distinct peaks were observed and corresponded to higher oligomers or aggregates and monomeric forms of Erb1. Recombinant protein resulted to be highly soluble and the fractions of Erb1 eluted from Hitrap Heparin column could be concentrated up to 50mg/ml.

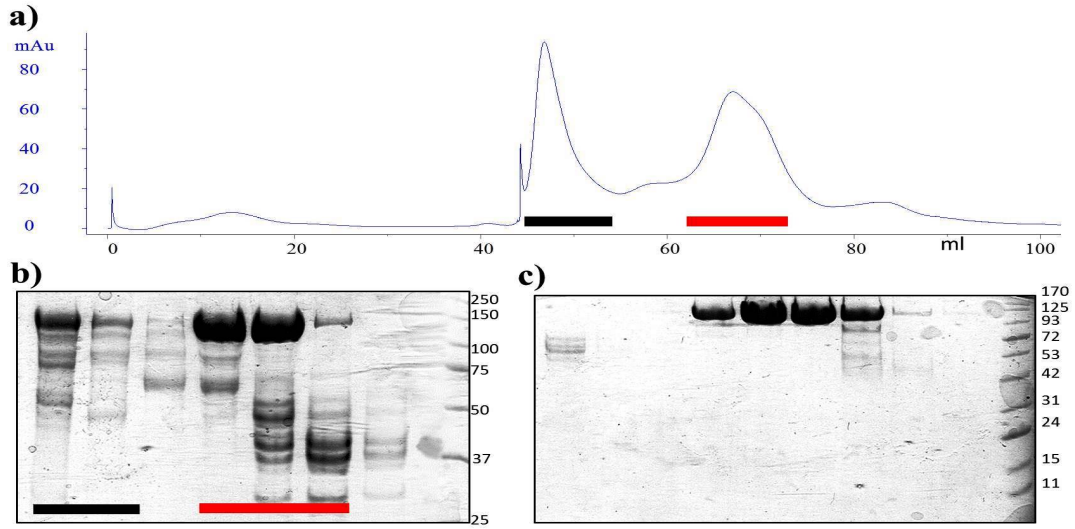


Figure 4.1 Purification of Erb1.
a) Chromatogram showing elution profile from gel filtration column. The monomer of Erb1 eluted at 68ml, volume equivalent to 135 kDa b) 10% SDS-PAGE shows fractions corresponding to the chromatogram. Black bar indicates fractions corresponding to higher oligomerization states or aggregates of Erb1. Red bar marks fractions containing monomeric Erb1 that were then passed through HiTrap Heparin in order to remove contaminations. c) 10% SDS-PAGE representing fractions eluted from HiTrap Heparin column. Molecular weight of each band in the MW marker is shown in kDa.

4.1.2. Expression and purification of Nop7

6xHis-Nop7 (72 kDa) expression was achieved using E. coli (DE3) BL21 CodonPlus in LB medium induced with IPTG. The protein was purified combining IMAC chromatography (HisTrap) and gel filtration (Fig. 4.2). In size exclusion the protein eluted as a single, monomeric peak and was concentrated up to 45 mg/ml.

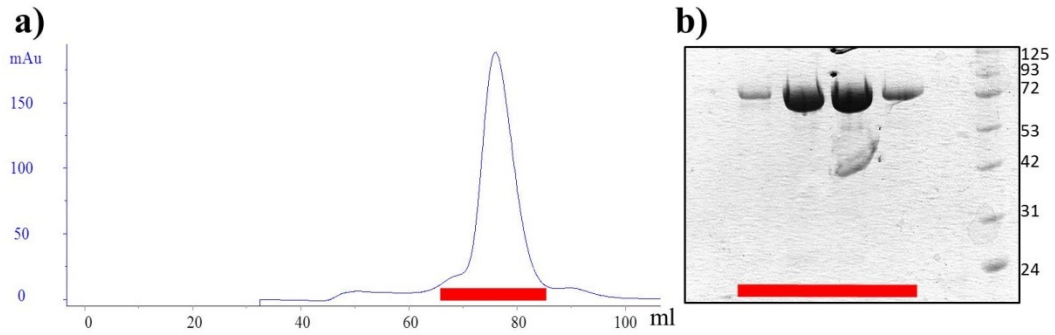


Figure 4.2 Purification of Nop7.
a) Chromatogram showing elution profile from gel filtration column. The monomer of Nop7 eluted at 77ml, volume equivalent to 85 kDa b) 10% SDS-PAGE shows fractions corresponding to the chromatogram. Red bar marks fractions containing monomeric Nop7

4.1.3. Expression and purification of Ytm1

Ytm1 could not be expressed in a soluble form in *E. coli*. Different expression condition, strains and media were tested but 6xHis-tagged protein remained in insoluble fractions. Re-folding protocol showed that high urea concentration (4M) was required to solubilize the inclusion bodies of 6xHis-Ytm1. Neither co-expression nor co-solubilization with Erb1, binding partner of Ytm1, yielded soluble protein. MBP-Ytm1 was soluble when expressed in bacteria but protein aggregation was detected in gel filtration probably due to improper folding. In a high-throughput screening for optimal conditions for protein expression at OPPF facility it was seen that soluble Ytm1 could be expressed with 6xHis-tag using Eukaryotic expression systems (Fig. 4.3). Small-scale transient expression in HEK cells and baculovirus-induced over-expression in Sf9 cells yielded soluble, non-aggregated protein. It was also seen that longer over-expression time (more than 72h) resulted in Ytm1 degradation.

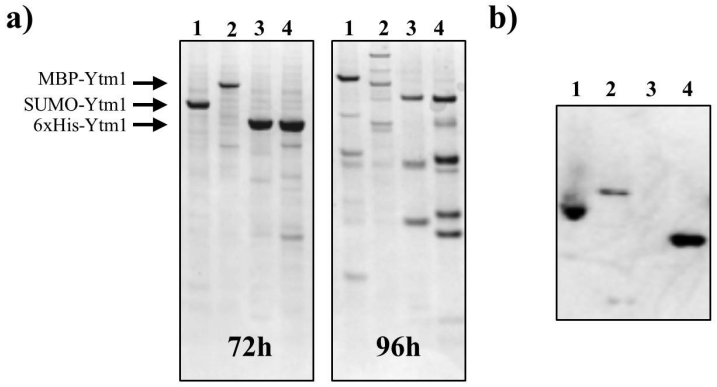


Figure 4.3 Small-scale expression tests of Ytm1.
a) SDS-PAGE of soluble protein extract. Sf9 insect cells were infected with 5 µl of P_0 stock of baculovirus carrying Ytm1 derived from 1: pOPINS3C 2:pOPINM 3:pOPINF 4:pOPINE. Soluble protein expression was checked after 72h (left) or 96h (right) upon infection. After 96h of expression severe protein degradation was detected. b) Western blot of transient expression in HEK293T cells using 1: Ytm1-pOPINS3C 2: Ytm1-pOPINM 3: Ytm1-pOPINF 4: Ytm1-pOPINE. Overexpressed proteins were detected with antiHis antibody.

After scaling up, 6xHis-Ytm1 (54kDa) was expressed in Sf9 cells, the protein was purified using IMAC chromatography and size exclusion column (Fig. 4.4). It eluted as a monomer and was concentrated up to 30 mg/ml.

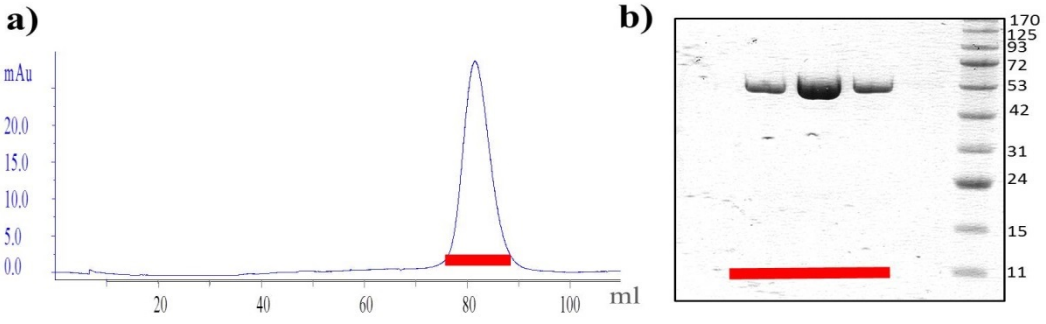


Figure 4.4 Purification of Ytm1.
a) Chromatogram showing elution profile from gel filtration column. The monomer of Ytm1 eluted at 82ml, volume equivalent to 55 kDa b) 10% SDS-PAGE shows fractions corresponding to the chromatogram. Red bar marks fractions containing monomeric Ytm1.

4.1.4. Complex formation and co-purification

Different approaches have been made in order to obtain pure, homogeneous ternary complex. Initially, Nop7 and Erb1 were purified separately, mixed in equimolar amounts and resulted to co-elute together in gel filtration giving a stable Nop7-Erb1 complex that was used for crystallization trials. Upon successful expression of soluble Ytm1 in Sf9 cells, the complete trimer was seen to hold together in size exclusion chromatography (Fig. 4.5).

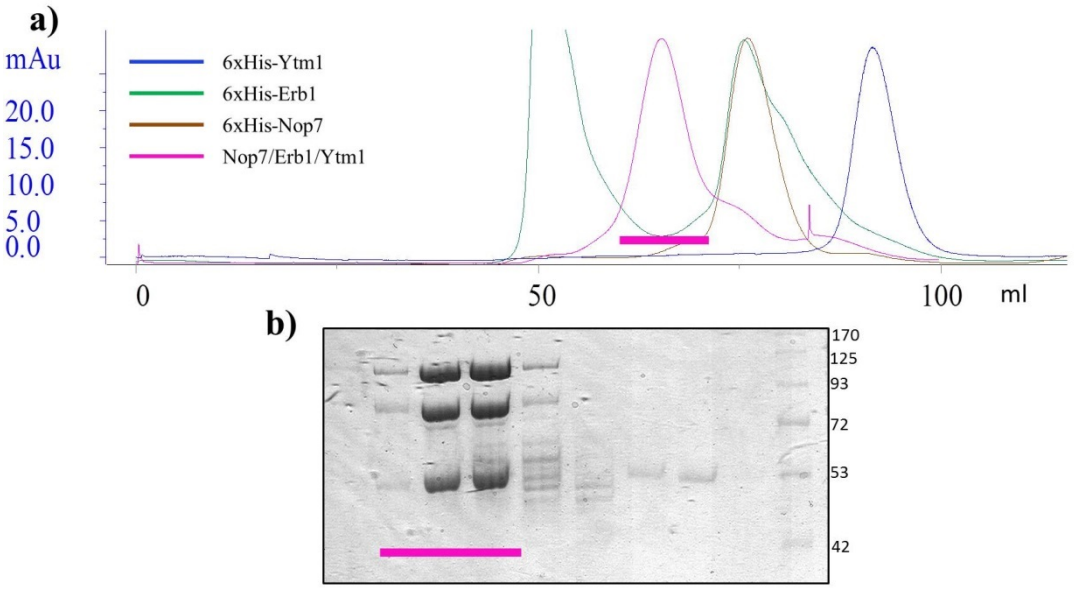


Figure 4.5 Reconstitution of Nop7-Erb1-Ytm1 complex.
a) Chromatograms showing elution profiles of the three proteins separately (Ytm1 in blue, Nop7 in brown and Erb1 in green) compared to the one of the trimer (pink line). Noticeable shift in elution volume indicated a stable heterotrimer formed by the proteins. Elution volume of 61ml was equivalent to 320 kDa b) 10% SDS-PAGE of the fractions corresponding to the chromatogram of Nop7-Erb1-Ytm1 confirmed that the proteins co-eluted together. Pink bar marks fractions that were pooled and concentrated up 30 mg/ml.

Additionally, in order to prevent protein degradation from the very initial stages of purification, co-purification was performed and consisted in simultaneous lysis of the cells that had over-expressed Erb1, Nop7 and Ytm1 separately. This robust approach did not allow to exactly determine the initial amount of each protein

in the cell pellets that were mixed together. Nevertheless, the HisTrap column used in the first step purification retained with higher affinity the trimer because Erb1, Nop7 and Ytm1 had 6xHis-tag. Finally, any excess of free-protein could be removed in gel filtration due to the difference in size between the ternary complex (220 kDa) and monomeric species of the proteins, however higher oligomerization states were also detectable (Fig 4.6). Pure Nop7-Erb1-Ytm1 complex was then used for crystallization assays.

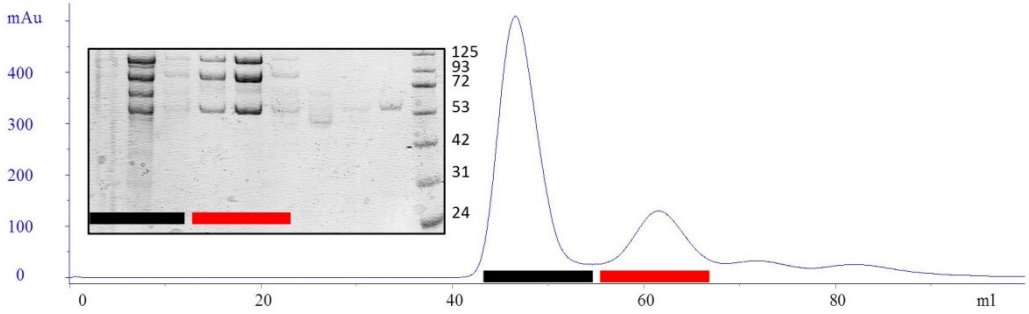


Figure 4.6 Co-purification of Nop7-Erb1-Ytm1 complex.
Chromatogram and SDS-PAGE of Nop7-Erb1-Ytm1 that was previously co-purified on HisTrap column. First peak (black bar) corresponds to a higher MW species and its intense absorbance at 280nm could indicate presence of nucleic acid in the sample. The second peak (red bar) eluted at 61ml (325 kDa) indicating that a stable heterotrimer was formed.

4.2. Crystal structure of Erb1$_{416-807}$ from *Saccharomyces cerevisiae:*

4.2.1. Crystallization and crystal optimization

Stable dimer formed by Erb1 and Nop7 was used in a set of crystallization screenings that after 4 to 5 days yielded diamond-shaped small crystals (Fig. 4.7).

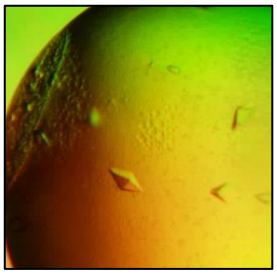


Figure 4.7 Crystallization trials of Nop7-Erb1.
Diamond-shaped crystals grew in 0.1M Hepes pH 7.5, 8% ethylene glycol and 10% PEG 8000 and diffracted poorly to 2.9Å

Those crystals diffracted poorly at 2.9Å and the collected dataset did not allow us to obtain phases. Nevertheless, unit cell dimensions already suggested that the entire dimer would not fit into asymmetric unit.

Table 4.1 Crystals grown with additives.
A set of different additives were added to the mother liquor from the original crystal screen. Crystals appeared in several conditions and those that grew with addition of ethanol diffracted to the highest resolution

Additive	Concentration	Resolution (Å)
Taurine	0.1 M	5.4
γ-Butyrolactone	40% v/v	2.5
Ethanol	30% v/v	1.6
Methanol	30% v/v	1.8
1-Butanol	7% v/v	2.04
1,4-Dioxane	30% v/v	2.09
Dichloromethane	0.75% v/v	5.2

In order to improve diffraction quality an optimization screening with additives was set up and gave rise to a number of similar crystals. Table 4.1 summarizes obtained diffraction depending on the additive used in the screen. Crystals that grew with addition of ethanol diffracted up to 1.6 Å and the collected dataset was used for phasing. Data collection statistics are shown in Table 4.2.

Erb1$_{416-807}$	
Wavelength (Å)	0.9763
Resolution range (Å) [a]	58.07 - 1.6 (1.657 - 1.6)
Space group	P 21 21 21
Unit cell	
a, b, c (Å)	52.02, 62.43, 158.22
α, β, γ (°)	90, 90, 90
Total reflections	446057 (43775)
Unique reflections	67582 (6545)
Multiplicity	6.6 (6.7)
Completeness (%)	98.07 (96.78)
Mean I/sigma(I) [b]	19.48 (2.16)
Wilson B-factor	22.85
R-merge	0.04807 (0.8424)
R-meas[c]	0.0523
CC1/2	0.999 (0.834)
CC*	1 (0.954)

Table 4.2 Data collection statistics of Erb1$_{416-807}$.
[a] Statistics for the highest-resolution shell are shown in parentheses.
[b] Mean $[I/s(I)]$ is the average of the relation between the intensity of the diffraction and the background.
[c] $R_{meas} = \{\Sigma_{hkl} [N/(N-1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle|\} / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and N is the multiplicity of reflection *hkl*.

4.2.2. Crystallization of Nop7-Erb1 dimer yields crystals of Erb1$_{416-807}$ due to sever degradation of Erb1

Since the volume of the unit cell was not high enough to accommodate the dimer of Nop7 and Erb1 (128442.9 Å^3) we asked if any fragment of those proteins could have been crystallized. In order to predict the potential degradation of macromolecules we assessed the stability of Nop7, Erb1 and Nop7-Erb1 dimer incubating the samples at 4ºC and we observed that the concentrated sample of Erb1 and Erb1/Nop7 presented a pattern of sever degradation corresponding to Erb1 (Fig. 4.8). In order to understand the nature of the degradation we analyzed the sequence of the protein and confirmed the presence of a PEST motif that precedes the sequence of the β-propeller. PEST regions are often flexible, exposed and easily accessible by proteases and are implicated in a fast turnover of eukaryotic proteins.

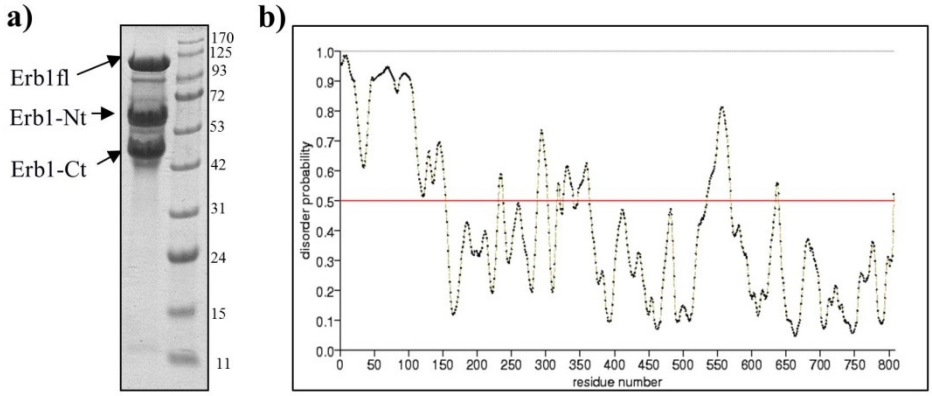


Figure 4.8 Erb1 suffers severe proteolytic degradation.
a) SDS-PAGE of a sample containing Erb1 that was left over-night at 4ºC shows that approximately 50% of the protein cleaved into two peptides that were identified by MALDI-TOF as N-terminal and β-propeller domains of Erb1. b) Prediction of protein disorder indicates that the N-terminal end of Erb1, the central region around residues 330-360 and the flexible segment within β-propeller domain exceed the threshold (red line) suggesting possible areas prone to suffer proteolysis.

Because the C-terminal domain of Erb1 was predicted to contain 7 WD repeats that would form a β-propeller domain we performed molecular replacement using models of other seven-bladed β-propellers. Successful phasing confirmed that the

crystals only contained the carboxy-terminal fragment of Erb1. Obtained phases allowed us to trace good quality model of the propeller (Table 4.3 and Fig 4.9).

Erb1$_{416-807}$	
R-work	0.1602 (0.2678)[a]
R-free	0.1740 (0.2891)[b]
Number of non-hydrogen atoms	3302
macromolecules	2928
ligands[c]	17
water	357
Protein residues	356
RMSD[d] (bonds)	0.006
RMSD[d] (angles)	1.07
Ramachandran favored (%)	97
Ramachandran outliers (%)	0
Average B-factor (Å)	31.10
Macromolecules (Å)	30.00
Ligands (Å)[c]	44.60
Solvent (Å)	39.10

Table 4.3 Refinement statistics of Erb1$_{416-807}$.
[a] R-work = S_{hkl} {[$F_{obs}(hkl)$] - [$F_{calc}(hkl)$]} / S_{hkl} [$F_{obs}(hkl)$], where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.
[b] R-free corresponds to R_{factor} calculated using 2 % of the total reflections selected randomly and excluded during refinement.
[c] Ligands: glycerol, ethylene glycol, ethanol
[d] RMSD is the root mean square deviation

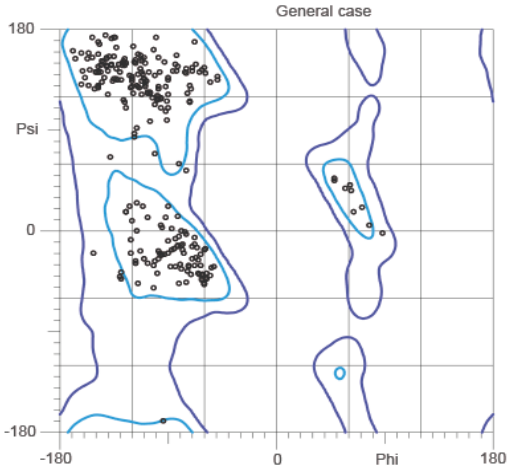


Figure 4.9 Ramachandran plot of Erb1$_{432-807}$ model.
97% of residues were found to fall into favored region whereas 3% was in allowed regions. No Ramachandran outliers were seen in the model.

4.2.3. Structural details of Erb1$_{416-807}$

The crystal structure of Erb1Ct confirmed that it folded into a seven-bladed β-propeller. The seventh blade is formed by three β-strands from the last WD repeat and one strand from the first WD repeat forming so-called "velcro" closure of the propeller (Fig. 4.10). In general, the domain superposes well with other β-propellers formed by seven WD repeats. Structural analysis done with Dali Server showed that over 200 structures deposited in PDB superimposed with Erb1$_{416-807}$ with RMSD ≤ 2.5 Å. In spite of the generally well conserved architecture of WD40 domain, Erb1Ct presented a series of important differences that could be associated with its specific function.

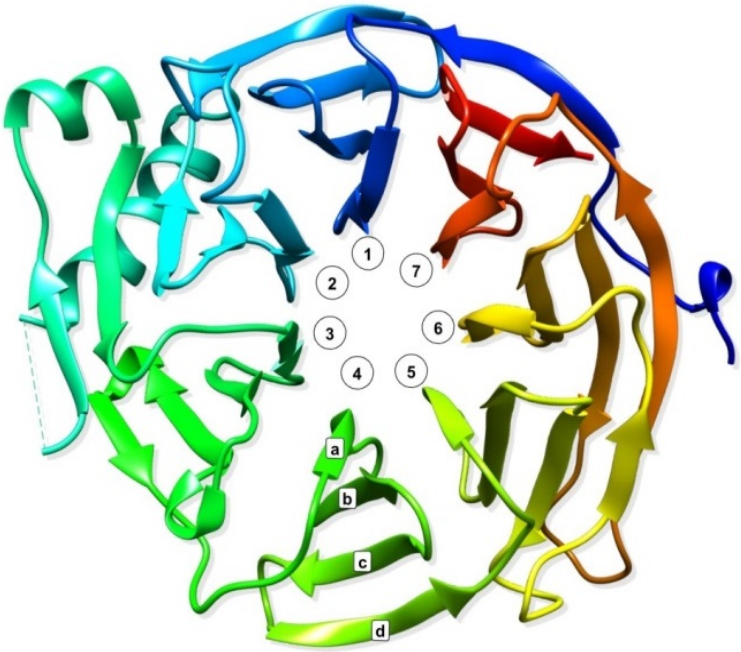


Figure 4.10 Cartoon representation of the β-propeller domain of Erb1.
Seven β-sheets form the circular shape of the propeller. Blade numbering is shown in the center; strand nomenclature is represented for blade 4. Velcro-like closure appears within blade 7 and comprises strand "d" from WD1 and strands "a", "b" and "c" from WD7.

The sequence of the domain is well conserved within Erb1/Bop1 family (Fig. 4.11) but is quite dissimilar to other β-propeller containing proteins. Curiously, none of the seven WD repeats harbors the eponymous WD di-peptide at the end of strand "c".

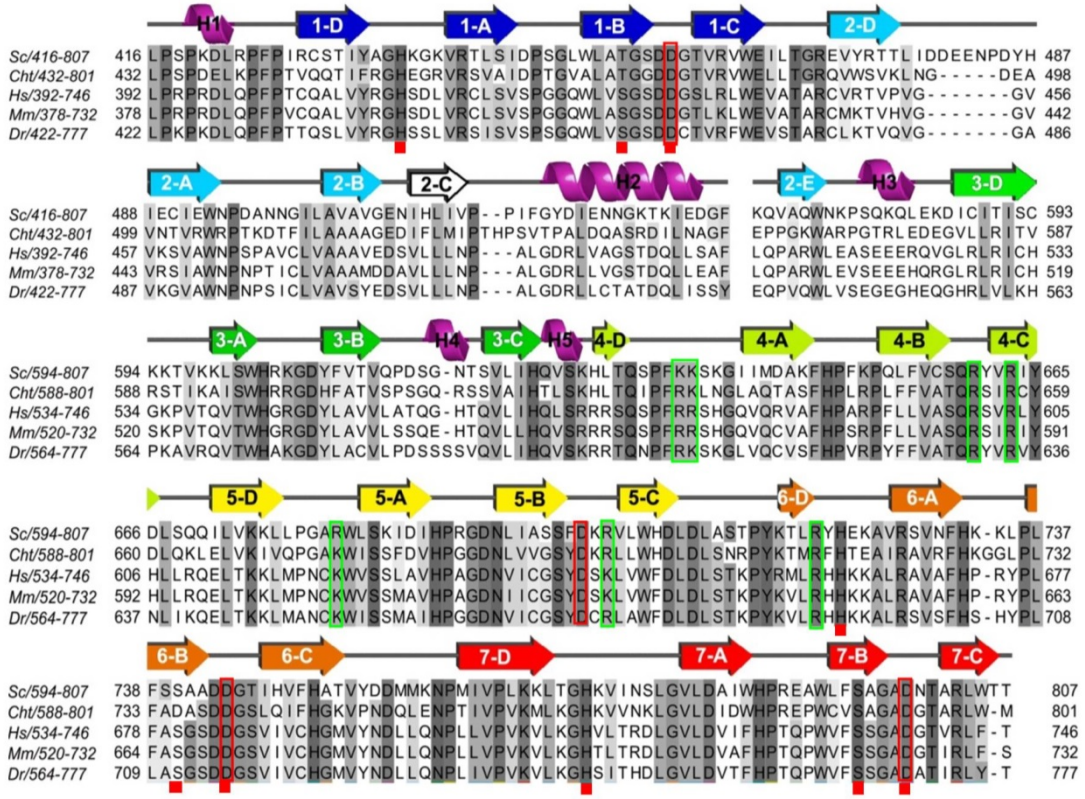


Figure 4.11 Sequence multi-alignment of the C-terminal domain of Erb1.
Sequences corresponding to the β-propeller domain of Erb1 have been aligned using T-coffee server. Sc: *Saccharomyces cerevisiae,* Cht: *Chaetomium thermophilum*, Hs: *Homo sapiens*, Mm: *Mus musculus*, Dr: *Danio rerio.* Conserved residues are depicted with shading. Elements of secondary structure elucidated from the model of $Erb1_{416\text{-}807}$ are shown over the alignment and are colored according to the WD repeats. Conserved amino acids that form part of the structural triad are marked with red squares. Red rectangles show aspartic acids from "b-c" loops. The green rectangles correspond to the conserved residues that form positively-charged pocket on the surface of the propeller

From an evolutionary point of view, the most conserved feature of WD40 repeats is the presence of an aspartic acid residue in a loop between strands "b" and "c" that is involved in structural triad formation. This triad is necessary for a proper organization of each blade and includes, in addition to aspartic acid, a histidine from "d-a" loop and a serine/threonine from strand "b" (Fig. 4.12). The three residues establish a network of electrostatic interactions that make the loop "d-a" more rigid thus stabilizing interblade organization.

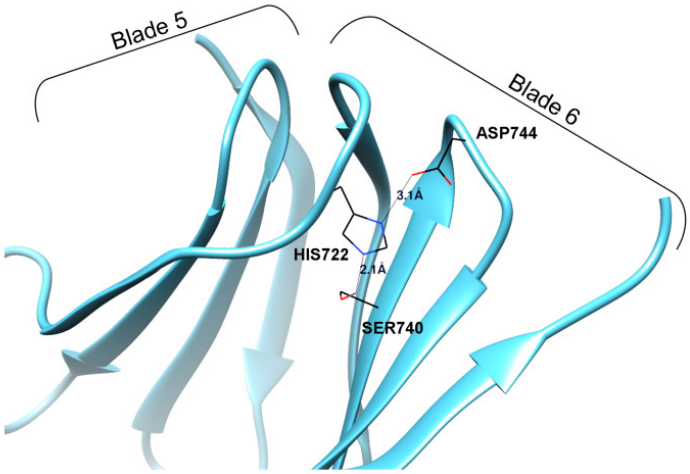


Figure 4.12 Structural triad present between blades 5 and 6 of Erb1$_{416-807}$.
An example of a canonical triad that guarantees proper blade organization shows hydrogen bonding between a histidine from "d-a" loop, a serine from strand "b" and an aspartic acid from "b-c" loop. The length of bonds is shown in Å.

In Erb1 five "b-c" loops contain an aspartic acid but only four of them are conserved (WD repeats 1, 5, 6 and 7) (red rectangles on Fig. 4.11). The canonical triad appears in blades 1, 6 and 7. "b-c" loop from WD2 contains a glutamic acid and in WD4 there is a glutamine, both residues are conserved and are involved in stabilization of the blades but the network of interactions is not similar to the one that takes place in the structural triad described previously. WD3 does have an aspartic acid in the "b-c" loop but it is not conserved and in human Bop1 there is a glutamine in its place. At last, in the WD5 repeat fully conserved aspartic acid does not form triad because the corresponding histidine is not present in the "d-a" loop.

4.2.4. Erb1 contains a long insertion within WD repeat 2

As described in Wegrecki et al., the most distinctive feature of the core of the β-propeller in this study is the blade 2 which due to an insertion contains five, and not four, β-strands and shows a protrusion attributed to two α-helices (H2 and H3 on Fig. 4.11). Electron density map allowed us to trace and build model for residues 515-534 and 571-594, being the rest of the insertion unmodeled. This missing part seems to be Fungi specific as it becomes much shorter in higher eukaryotes (Fig. 4.11). Helix H2 (residues Y520-D532) appears between strands "2c" and "2e" and is attached to the base of the propeller (Fig. 4.13a).

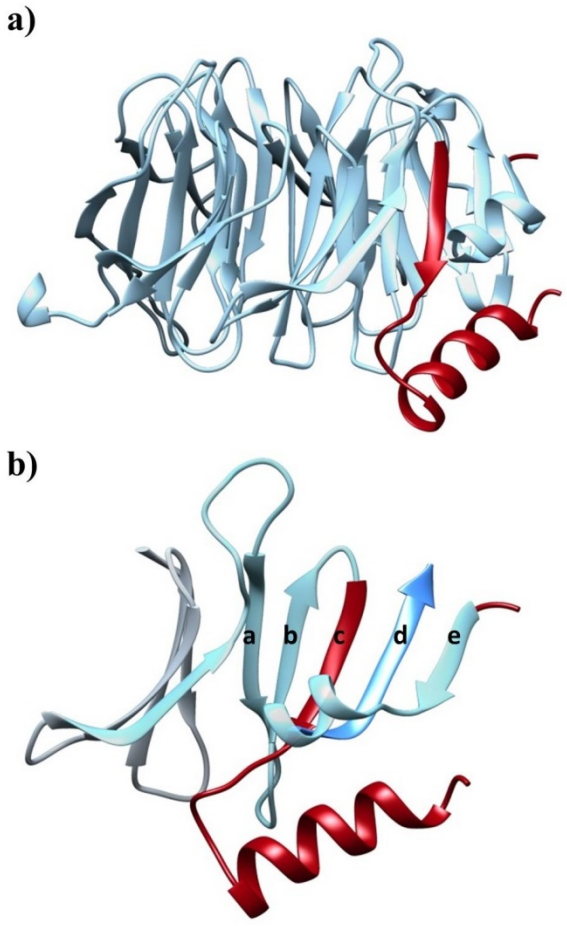


Figure 4.13 WD2 contains a long insertion.
a) The insertion (in red) forms additional β-strand and an α-helix that protrudes from the globular β-propeller. b) Position of the insertion (red) in the context of the second blade only. Residues corresponding to WD2 are represented in light blue and the strand "d" of WD3 is shown in dark blue

In general, the sequence of the helix is poorly conserved, but it contains two invariable non-polar residues: I522 that makes hydrophobic contacts with the backbone of "2a-2b" loop and I530 which interacts with the C-terminal fragment of strand "3d". The β-sheet corresponding to this blade is formed by strands "2a", "2b", "2c", "3d" and "2e" clearly indicating an alteration of a standard WD40 pattern (Fig. 4.13b). Whereas strand "3d" unambiguously indicates the beginning of WD3, the sequence of WD repeat 2 does not show any significant conservation but still contains strategic residues that allow formation of hydrophobic and electrostatic interactions with neighboring blades. Initial sequence-based analysis suggested that between WD repeats 2 and 3 there was an approximately 80-residue long segment which did not contain any WD pattern. Surprisingly, when we aligned the sequence of Ct domain of Erb1 with other non-Erb1/Bop1 β-propeller-containing proteins we could clearly see that W575 from strand "2e" corresponded to tryptophan residue from WD dipeptide that typically appears in strand C (as in human WDR5 protein, PDB: 2H14) (Fig. 4.14).

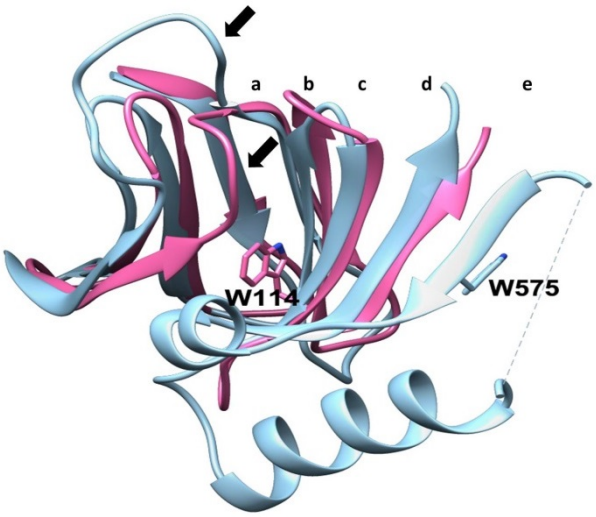


Figure 4.14 Conserved tryptophan is misplaced due to insertion.
Comparison of blades 1 and 2 of WDR5 from *H. sapiens* (PDB:4CY2; pink) and Erb1 (blue). Side chains of conserved tryptophan corresponding to the strand "c" (in canonical WD repeats) are shown for both proteins. Black arrows indicate the position of "2d-2a" loops

This fully conserved residue establishes important hydrophobic interactions with I592 from strand "2d" and H629 located in "3d" that are likely to be required for a proper conformation and attachment of the insertion to the side of the blade 2. We conclude that from an evolutionary point of view strand "e" corresponds to

strand “c” from a canonical blade although displaced, in the second blade of Erb1, by a segment containing “2c”-loop-“H2”. This insertion produced an important reorganization of the whole blade, altering the position and function of W-D dipeptide (W-N in this case). As a result, the second blade lacks the important tryptophan residue at the end of strand “2c” that would guarantee correct approach between blades 1 and 2. We observe that in this case there is a different interaction network, conserved in Erb1/Bop1 family but not in other WD repeat-containing proteins, that involves strand "2d” from blade 1 and a short α-helix, H3, from blade 2 (Q580-K585). This helix inserts between strands “2e” and “3d” and possesses two non-conserved lysine residues (K581; K585) that interact with loop “2d-2a” through hydrogen bonds. In consequence of this arrangement, α-helix H3 forms a lid that orientates close to a very hydrophobic area in blade 2 created by a segment of well conserved polar residues from strand “2b”. It is important to keep in mind that loop “2d-2a” is quite flexible and its vertical orientation makes the whole interface between blades 1 and 2 more opened when compared with the gaps between other blades which are completely covered by “d-a” loops (black arrows in Fig. 4.14).

4.2.5. Conserved residues form putative ligand binding areas on the surface

The best known and most obvious function of β-propellers is their capacity to establish multiple interactions with other proteins. The intrinsic rigidity and the shape of the domain create three well defined zones where the binding partner can attach: the top, the bottom and the circumference of the propeller. We searched for conserved residues on the surface of Erb1Ct which could indicate a region important to preserve protein interaction interface. There is a very clear division between a poorly conserved area, that includes blades 1, 2 and the upper part of blade 3, and a less variable surface of blades 4, 5, 6 and 7. In the bottom part, between blades 3 and 4 we identified a well conserved pocket which is a good candidate as a possible place of association with a binding partner.

In addition, it has been proposed that the central channel of WD domains could work as a scaffold that adapts for recognition of different ligands through side-chains of three residues from each blade: the one right before the strand “a” (A-1), the one just after strand “b” (B+1) and the second residue in the strand “a” (A2), thus making this portion of the propeller an universal but variable binding motif.

When we inspected these positions in Erb1Ct a strong conservation, especially in blades 1, 5, 6 and 7, was observed. Our findings were confirmed by WDSP web server which predicted hot-spot residues on the surface that were likely to be responsible for high-affinity interactions with other proteins (Fig. 4.15a). Moreover, those conserved positions seem to be related to Erb1/Bop1 function because they vary when compared to seven-bladed propellers from other families. Nevertheless, three of these superficial conserved amino acids, D457, R727 and D743 are also invariable in another family of WD repeat-containing proteins called Lis1 where they were shown to be involved in recognition of other macromolecules (Fig. 4.15b).

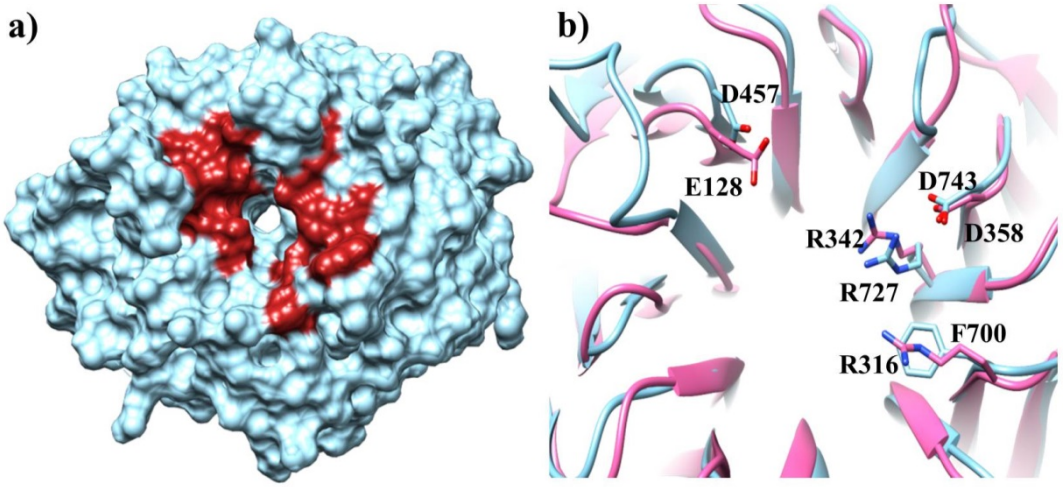


Figure 4.15 The top entrance of the β-propeller contains "hot spot" residues.
(a) Top face of the propeller showing the position (in red) of the residues most likely to participate in macromolecular interactions as predicted by WDSP server (http://wu.scbb.pkusz.edu.cn/wdsp/). (b) Superposition of Lis1 (pink, PDB: 1VYH) with Erb1 β-propeller. The side chains of the conserved positions are shown and labeled

4.2.6. The surface of $Erb1_{416-807}$ is positively charged

Calculated theoretical isoelectric point (pI) of the β-propeller was 8.7 meaning that the protein was slightly alkaline. When we analyzed the electrostatic potential of the domain, a large stretch of positively charged residues could be visible on the major part of the surface (Fig. 4.16). We also used the PatchFinder Plus algorithm to identify the biggest positive patch on the surface of the domain.

Indeed, as seen for the electrostatic surface analysis, the tool found a big region of basic residues on the surface that included five blades and the entrance to the central channel on the top face of the propeller. Those amino acids were predominantly located on one side of the propeller. We checked the conservation of the residues and we observed that seven of them were invariable in higher Eukaryotes (green rectangles in Fig. 4.11), thus the large positively charged area could have a functional relevance.

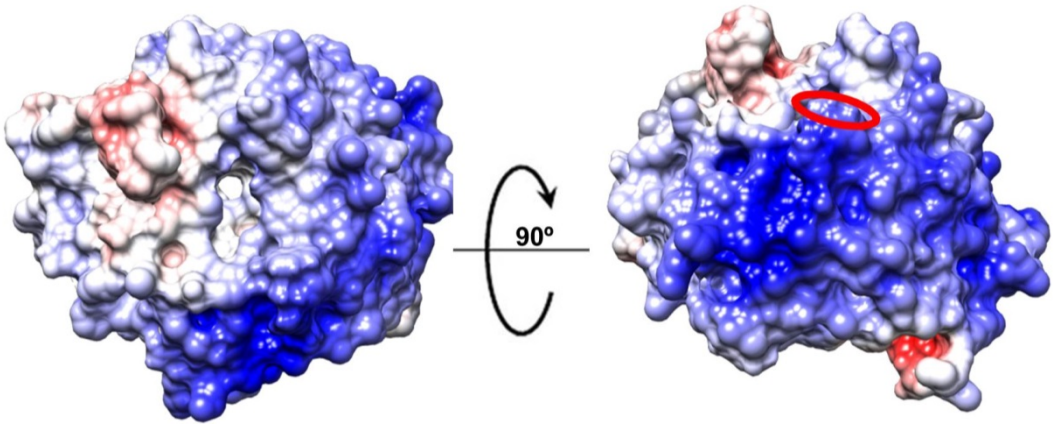
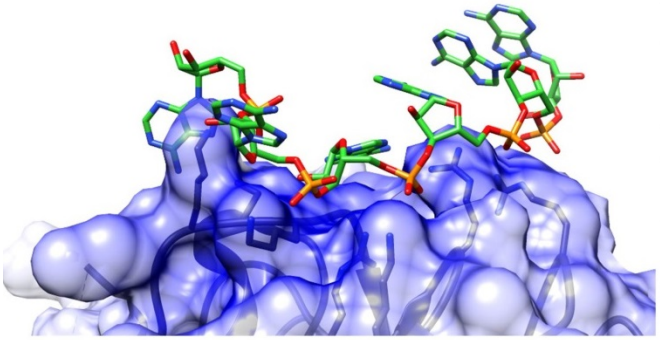


Figure 4.16 Surface of Erb1 β-propeller is positively charged.
Surface representation of the electrostatic potential of the domain (from red (-10) to blue (+10)kb T ec-1) as calculated with APBS. The top face is shown on the left and the most positively charged area formed by blades 4 and 5 is visible on the right panel. The red oval indicates the position of Trp682.

Because previous studies in yeast have shown that full-length Erb1 was able to bind rRNA within pre-ribosomal particle, we decided to assess whether the carboxy-terminal domain of Erb1 could be involved in the binding. We used HADDOCK server (http://haddock.science.uu.nl/index.html) to dock a 6-nucleotides long fragment of RNA and we could observe that the single stranded poly(A) was nicely fitted into the positively charged area (Fig. 4.17). In order to empirically test if the protein could bind nucleic acids we used the β-propeller domain of Erb1 from *Chaetomium thermophilum* due to its enhanced expression levels and stability as described later. Detailed results of RNA bidning experiments are described in section 4.4.



HADDOCK score	-79.9 +/- 6.3
Cluster size	29
RMSD from the overall lowest-energy structure	1.1 +/- 0.8
Van der Waals energy	-24.0 +/- 6.3
Electrostatic energy	-410.6 +/- 50.0
Desolvation energy	24.7 +/- 8.4
Restraints violation energy	14.8 +/- 24.29
Buried Surface Area	731.2 +/- 129.8
Z-Score	-1.6

Figure 4.17 Six-nucleotide long poly(A) was docked into electropositive patch of Erb1$_{416-807}$. HADDOCK server was used to perform docking of RNA and the β-propeller and proved that ssRNA molecule can fit into the area formed by conserved positively charged residues (shown in blue) on the surface of Erb1$_{416-807}$. Docking results for the best cluster are shown on the right.

4.3. Reconstitution of Nop7-Erb1-Ytm1 complex from *Chaetomium thermophilum*

4.3.1. Generation of cDNA library of *Chaetomium thermophilum*

Since the availability of a complete genome of a thermofilic fungus has been considered promising for structural studies of eukaryotic proteins (Amlacher et al., 2011), an attempt to reconstitute Nop7-Erb1-Ytm1 complex from *C. termophilum* has been made. We decided to generate the cDNA from the transcriptome of the fungus because many of its genes, unlike in *S. cerevisiae*, contain introns. The organism was grown in laboratory conditions until a considerable amount of biomass was reached and was sufficient for the extraction of RNA. Agarose-gel electrophoresis showed that the quality of the material was satisfactory to perform reverse transcription (Fig. 4.18a).

Resulting cDNA was then used as a template for several PCR reactions in order to amplify genes of interest. Successful amplification was achieved for *erb1* and *ytm1* genes but not for the full ORF of *nop7* gene (Fig. 4.18b). Nevertheless fragments of the gene were amplified and allowed expression of certain domains of ChNop7 protein. Because *nop7* gene was predicted not to contain introns, whole genomic DNA extraction was performed and resulting DNA was successfully used

to amplify *nop7* ORF. The genes were subsequently cloned into expression vectors as described in section 3.1.2.

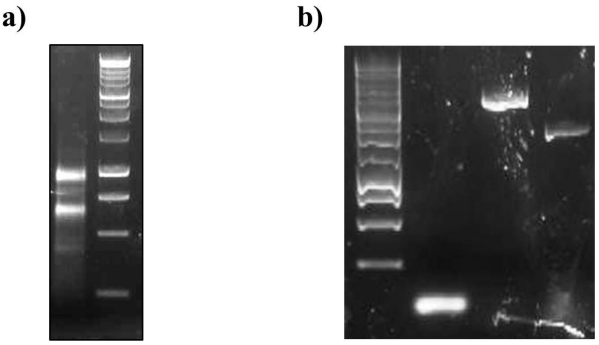
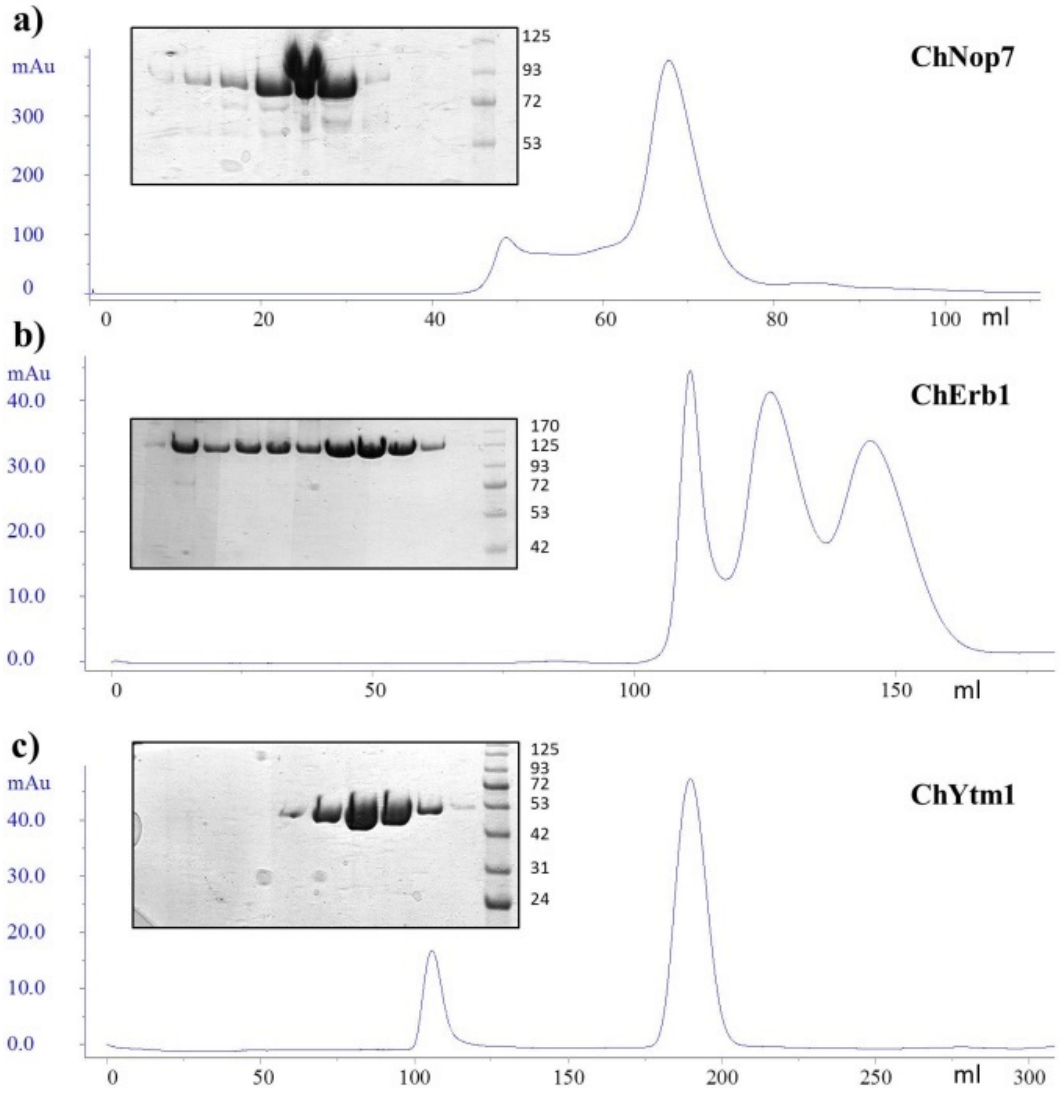


Figure 4.18 RNA extraction from *C. thermophilum.*
a) 3 µl of extracted sample were loaded onto agarose gel to assess the quality of the obtained RNA. b) cDNA library was used to amplify *erb1* (lane 3) and *ytm1* (lane 4) genes. *nop7* gene (lane 2) could not be obtained in the same way. DNA ladder is shown in lane 2 of panel a) and lane 1 of panel b)

4.3.2. Expression and purification of ChErb1, ChNop7 and ChYtm1.

The full-length proteins were over-expressed in the same manner as their counterparts from yeast although the yield was higher in each case. Proteins were purified according to the protocols optimized for Nop7, Erb1 and Ytm1 from *S. cerevisiae.* 6xHis-ChNop7 (80kDa) and 6xHis-ChYtm1 (55.5 kDa) were highly monodisperse and appeared as monomers in gel filtration whereas 6xHis-ChErb1 (93 kDa), as in case of Erb1 from yeast, resulted to form monomers, dimers and larger oligomers (Fig. 4.19). The multimeric state of ChErb1 was also checked by native gel electrophoresis as described later.

Figure 4.19 Final results of purifications of full-length proteins from *C. thermophilum.*
a) Elution profile and corresponding SDS-PAGE of ChNop7 in Superdex 200 16/60 column. Mild protein degradation is visible. b) Elution profile and corresponding SDS-PAGE of ChErb1 in Superdex 200 26/60 column. Peaks corresponding to monomer, dimer and larger oligomer are visible in the chromatogram. c) Elution profile and corresponding SDS-PAGE of ChYtm1 in Superdex 200 26/60 column. Highly homogenous and pure sample was obtained



4.3.3. Proteins from *Chaetomium thermophilum* are more thermo-stable

In order to verify if proteins from a thermophile would present higher degree of stability we performed thermal shift assay (Thermofluor) that allowed us to measure the melting temperature (Tm) of a given protein in different chemical conditions. Then, we compared resulting Tm with that calculated for proteins from *S. cerevisiae*.

As seen for Nop7 and ChNop7, the ortholog from *C. termophilum* showed an overall higher thermal stability. In the conditions tested, from those that resulted in a melting curve suitable for Tm calculations, ChNop7 was more stable than Nop7 in 100% of cases (Fig 4.20a).Similar effect of enhanced thermal stability was observed when we analyzed Ytm1/ChYtm1. Again, the protein from *Chaetomium thermophilum* unfolded at higher temperatures. Only in presence of 0.1M sucrose and 0.5M urea, Ytm1 from yeast resulted to be more stable than its counterpart from the thermophile (Fig. 4.20b).

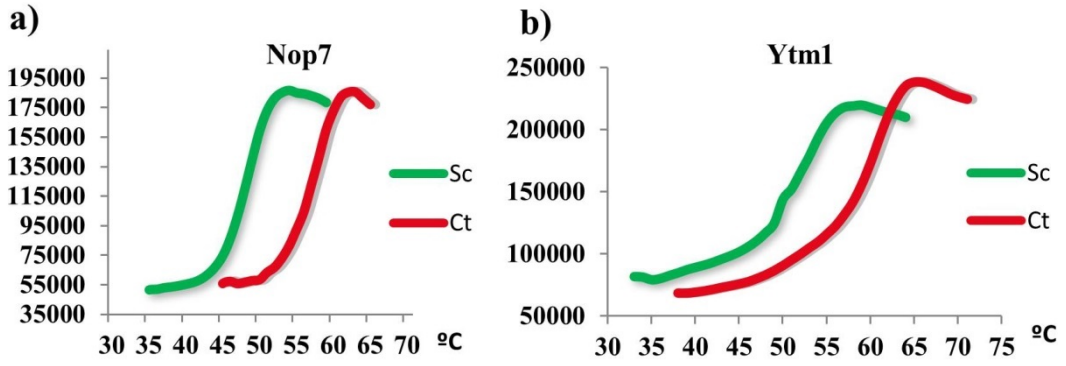


Figure 4.20 Proteins from *C. thermophilum* are more thermostable.
Thermofluor assay was performed in order to measure the melting temperature of the proteins in Buffer SE (50mM Hepes pH7.5, 150mM NaCl, 5% glycerol, 2mM β-mercaptoethanol). Increase in fluorescence was measured for Nop7/ChNop7 (a) and Ytm1/ChYtm1 (b). The proteins from yeast are depicted in green whereas their thermophilic counterparts are shown red.

4.3.4. Reconstitution of ChNop7-ChErb1-ChYtm1 complex in vitro

Initially we were able to express and reconstitute the dimer formed by ChErb1 and ChYtm1. The proteins were purified separately, mixed in equimolar amounts and co-eluted from gel filtration confirming their stable association. Upon successful cloning of ChNop7, we performed a co-purification and we obtained ChNop7-ChErb1-ChYtm1 trimer (220 kDa) that was used in crystallization trials. In order to completely remove nucleic acids that could have been co-purified with the complex, we included a HiTrap Heparin column step in purification. Size exclusion chromatography indicated that approximately 50% of the complex was in a higher

oligomerization state (Fig. 4.21a and b). The stoichiometry of the oligomer was difficult to estimate because the separation range of Superdex 200, resin we used for gel filtration, is not optimal for particles with such a high MW.

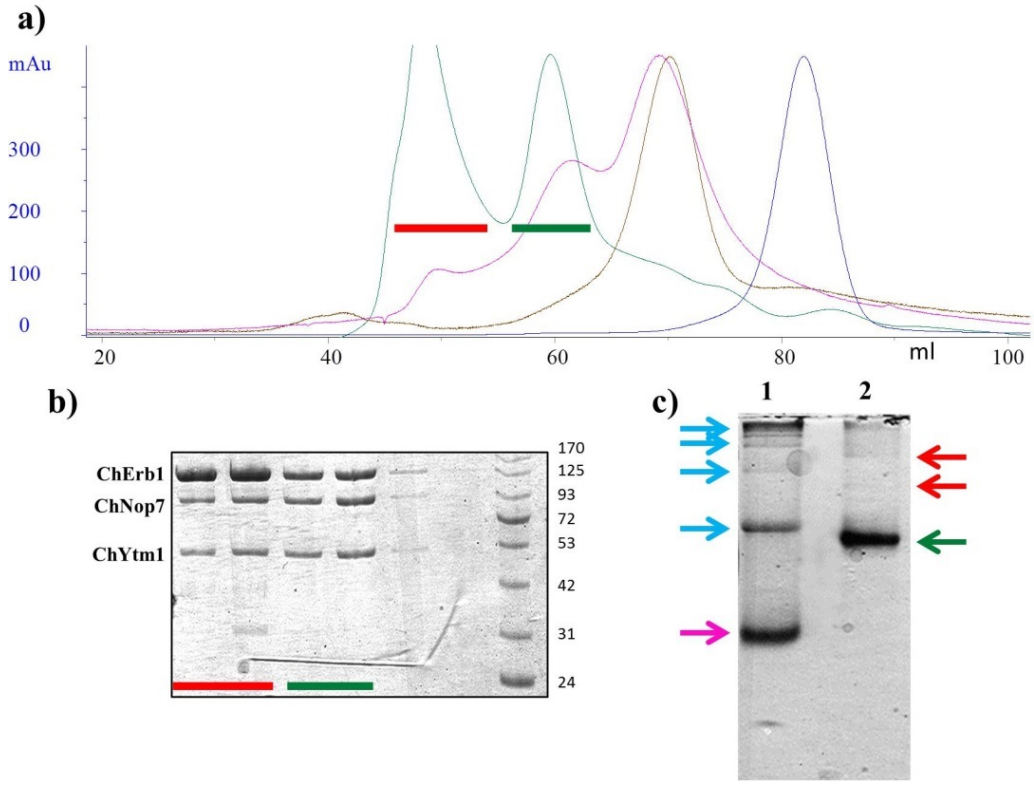


Figure 4.21 Stable trimer is formed by ChNop7, ChErb1 and ChYtm1 in gel filtartion.
Elution profiles in size exclusion chromatography (Superdex 200 16/60 column) of ChYtm1 (blue), ChNop7 (brown) and ChErb1 (pink) were superimposed with the one of the trimer (green). b) SDS-PAGE of the peaks from the chromatogram of the trimer (green) is shown and indicates that both peaks contain the complex. The second peak corresponds to the trimer (elution volume of 61 ml is equal to 320 kDa approx., green bar); the first one contains larger oligomer and excess of ChErb1 (red bar). c) native gel electrophoresis was done to check the oligomerization state of ChErb1 and the complex in solution. As seen in lane 1, monomeric form of ChErb1 is predominant (pink arrow) but at least 4 different oligomers (blue arrows) are also present. Lane 2 shows that ChNop7-ChErb1-ChYtm1 appears mainly as heterotrimer (green arrow) although small amount of larger complexes are also detectable (red arrows).

We performed native gel electrophoresis to check if we could observe different oligomers (Lane 2 Fig.4.21c) and we detected only very small amount of complexes larger than the trimer. The presence of higher MW oligomers was much more prominent in case of ChErb1 alone (Lane 1 Fig. 4.21c)

4.3.5. The amino-terminal domain of ChNop7 binds ChErb1

Previous studies have suggested that, in yeast, Nop7-Erb1 dimer was formed by the amino-terminal segment of Nop7 and a central region of Erb1. Based on those observations we decided to design truncated version of Nop7 or ChNop7 that would still bind to Erb1 or ChErb1, respectively, in vitro but would lack disordered regions and coiled-coli domains predicted on their C-termini (Fig. 22a).

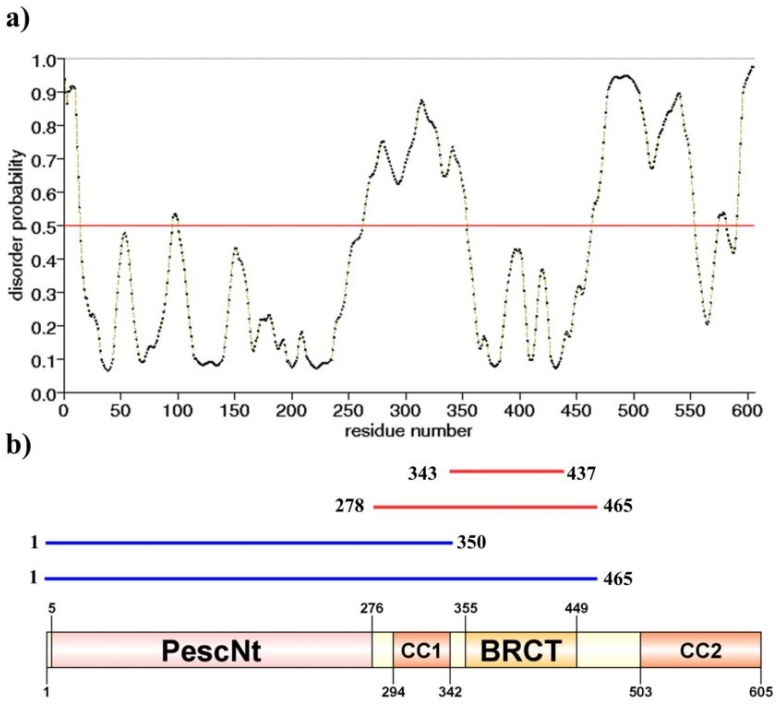


Figure 4.22 The C-terminal segment of Nop7 is disordered.
a) Disorder prediction plot for Nop7 from *S. cerevisisae* (identical plot was obtained for ChNop7) shows that only the most amino-terminal region and the BRCT domain are predicted to be ordered in solution. b) A scheme of domain organization of Nop7 and the truncated peptides that were assayed for its ability to bind Erb1. Blue bars indicate those fragment of the protein that efficiently interacted with Erb1 in gel filtration, whereas red bars depict those that did not (compare Fig. 4.23).

We were able to express and purify truncated versions of Nop7 and ChNop7: $Nop7_{1-465}$, $Nop7_{278-465}$, $ChNop7_{1-350}$ that contained the conserved N-terminal domain of unknown structure and $ChNop7_{350-466}$ that corresponded to the BRCT domain of ChNop7. All, except $ChNop7_{350-466}$, carried an amino-terminal 6xHis tag. The BRCT domain of ChNop7 was expressed as GST-fusion protein that was then cleaved and used without any tag for further experiments. Gel filtration confirmed that only $Nop7_{1-465}$ and $ChNop7_{1-350}$ were able to stably associate with Erb1 (Fig. 4.22b and 4.23). Because Holzel et al. observed that some point mutations within BRCT domain of Nop7 affected its binding to Erb1, we checked if BRCT domain of Nop7/ChNop7, although not essential for the interaction, could play role in Erb1/ChErb1 recognition. Using gel filtration we did not detect any binding between ChErb1 and $ChNop7_{350-466}$ (Fig. 4.23c).

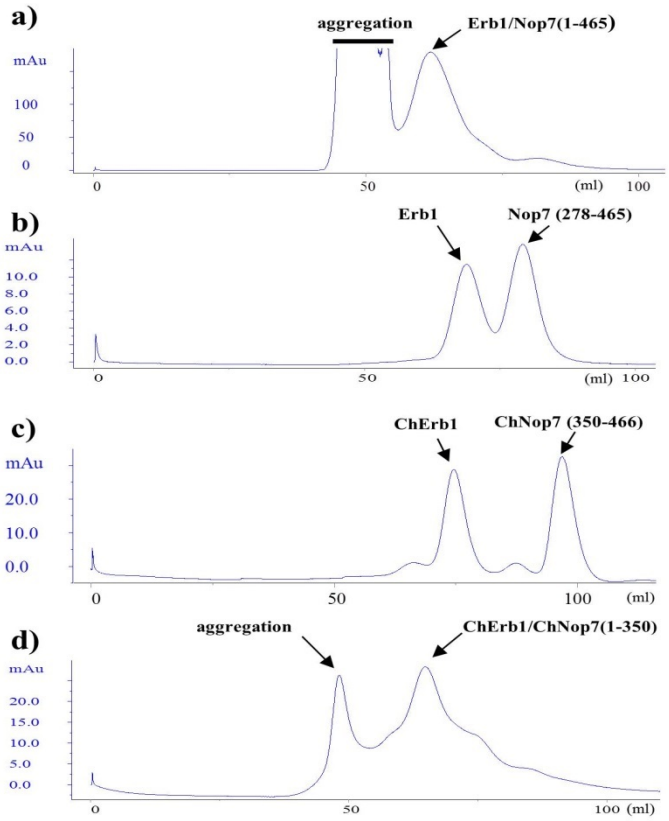


Figure 4.23 The amino-terminal domain of Nop7 is required for its interaction with Erb1.
Erb1 or ChErb1 were incubated with different truncated versions of Nop7 or ChNop7 and theirs stable association was then checked by gel filtration. Erb1 interacted with Nop7 that lacked its most C-terminal segment (residues 466-605) (a)), but stable binding was disrupted when the amino-terminal part was removed (residues 1-277) (b)). Whereas ChErb1 did not co-elute with the BRCT domain of ChNop7 (amino acids 350-466) (c)), it did associate with the N-terminal portion of ChNop7 (residues 1-350) (d)).

4.4. Carboxy-terminal domain of ChErb1 binds RNA in vitro

Our first attempts to perform in vitro binding assays failed because the C-terminal domain of Erb1 (residues 422-807) from yeast expressed poorly in *E. coli* and rapidly degraded during purification (Fig. 4.24a up). We decided to try whether the same domain from the thermophile would be stable enough to carry out the experiments. Since the sequence of the domain is well conserved between *S. cerevisiae* and *C. thermophilum*, including the basic residues from the putative RNA-binding area (shown with green boxes in Fig. 4.11), we considered ChErb1$_{432-801}$ to be suitable for validation of our findings based on Erb1$_{416-807}$ structure from yeast.

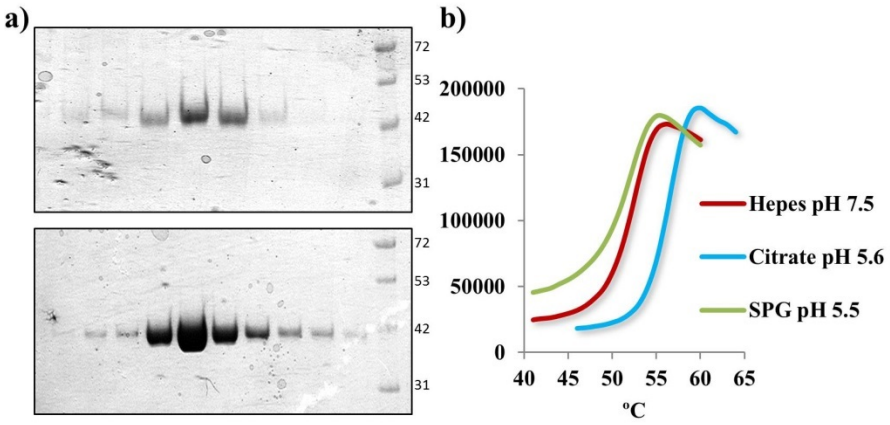


Figure 4.24 The β-propeller domain of Erb1/ChErb1.
a) SDS-PAGE shows fractions that contained Erb1$_{422-807}$ (top) or ChErb1$_{432-801}$ (bottom) that eluted from gel filtration column. Both proteins were purified from a 0.5L culture of *E. coli.* b) Thermal stability assays of ChErb1$_{432-801}$ showed that citrate at pH 5.6 resulted in higher Tm (blue line). This effect was produced not by the lower pH but rather by citrate itself because SPG buffer at similar pH did not enhance protein stability (green line)

ChErb1$_{432-801}$ expressed well in *E. coli* and large amount of protein eluted from HisTrap and gel filtration columns (Fig. 4.24a bottom). Nevertheless, the solubility of the molecule was not satisfactory as it tended to precipitate at concentrations higher than 5 mg/ml. Moreover, we observed that the protein was more stable when the maximum concentration did not exceed 1 mg/ml. Thermofluor showed that 0.1M citrate pH 5.6 produced an important shift in thermal stability of

the domain (Fig. 4.24b) but when the buffer composition was modified we did not observe any significant improvement in protein solubility. Finally we decided to maintain low concentration of the stock of ChErb1$_{432\text{-}801}$ because further experimental setup did not require concentrated samples.

Initially, we checked the affinity of the β-propeller of chErb1 for RNA in vitro using poly(U) agarose beads. As shown in the figure 4.25a the propeller appeared in the eluate from poly(U) beads. To investigate whether the interaction occurred through a well-defined surface that could be saturated, we incubated the protein with free poly(U) before it was loaded on the beads. The amount of the protein that could stably bind the poly(U) beads decreased in presence of 0.1mg/ml and 1mg/ml free poly(U), indicating saturation of the binding site (Fig. 4.25b).

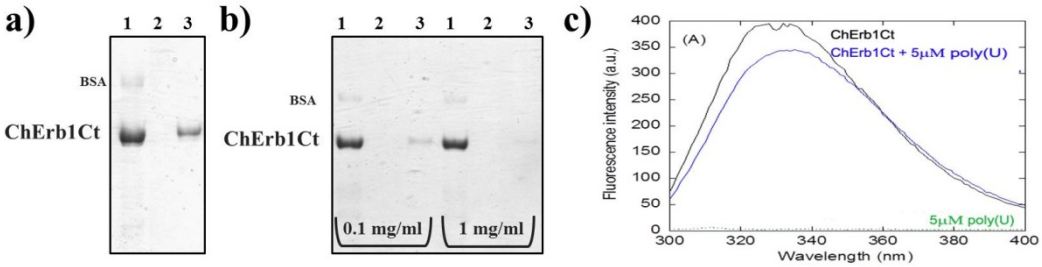


Figure 4.25 The β-propeller of ChErb1 binds nucleic acids in vitro.
a) ChErb1$_{432\text{-}801}$ was incubated with poly(U)-agarose beads (1:input) and after extensive washing (2:wash) it was detected in the elution (3) b) free polyuridilic acid (final concentration of 0.1 or 1 mg/ml) was added to the sample before it was loaded on the poly(U)-beads. Saturation of the binding site was detectable because lower amount of ChErb1$_{432\text{-}801}$ remained bound to the beads. 1:Input, 2:Wash, 3:Elution. c) Fluorescence emission spectra were acquired for ChErb1$_{432\text{-}801}$ alone (black) and in presence of 5 μM poly(U) (blue). Binding of RNA decreased the fluorescence intensity of ChErb1Ct. As a control 5 μM poly(U) alone was also measured (green).

Fluorescence experiments were carried out in order to confirm the binding between the protein and the nucleic acid. Emission spectrum of intact ChErb1$_{432\text{-}801}$ showed a maximum at 330 nm (black line in Fig. 4.25c). We acquired spectrum upon addition of 5 μM polyU (blue line in Fid. 4.25c), and we could observe a decrease in fluorescence intensity. In order to prove that the change in the spectrum of ChErb1 was due to the binding of RNA we showed that polyU alone did not

present emission upon excitation at 280nm (green line Fig. 4.25c). The fact that an exposed tryptophan (Trp682, red oval in the Fig. 4.16) is at the vicinity of the positively charged stretch (likely to participate in RNA binding), provides a good indication that the interaction takes place through the proposed area and could explain the change in fluorescence upon binding of the nucleic acid.

We next asked if the positively charged β-propeller domain would also bind DNA and using EMSA (Electrophoretic mobility shift assay) in native agarose gel we could detect binding between a 120bp long dsDNA and ChErb1$_{432\text{-}801}$ in a concentration-dependent manner (Fig. 4.26a).

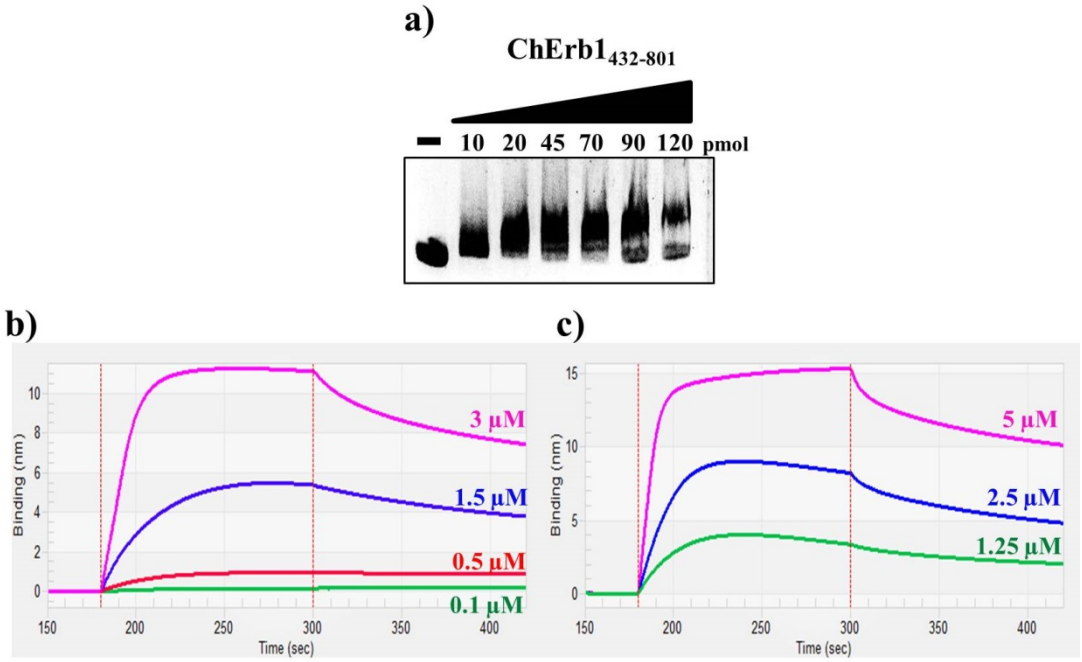


Figure 4.26. The β-propeller of ChErb1 binds preferentially to RNA.
a) agarose gel shows that upon addition of increasing amounts of the protein a clear shift in the electrophoretic mobility of a dsDNA (30ng) probe could be observed. In order to estimate bidning affinity Biolayer interferometry was performed: b) association and dissociation steps of different concentrations of ChErb1432-801 (3, 1.5, 0.5 and 0.1 µM) to a biosensor previously loaded with 15-nt long biotinylated poly(U). c) 20-nt long biotinylated DNA was loaded on a biosensor and then association of ChErb1432-801 (5, 2.5, 1.25 µM) was measured.

Finally, in order to precisely measure the binding affinity between ChErb1$_{432-801}$ and RNA or DNA we used biolayer interferometry. RNA binding was assessed by a 15 nucleotide-long biotinylated poly(U) immobilized on a Streptavidin Biosensor. The dissociation constant (K_D) was ~170 nM (Fig. 4.26b). In the second experiment, ssDNA fragment composed of 20 nucleotides carrying a biotin moiety on its 5' was used (Fig. 4.26c). Calculated K_D was ~400 nM and indicated that the C-terminal domain of ChErb1 bound to RNA with higher affinity than to DNA. Therefore, our in vitro binding assays have confirmed that the β-propeller domain binds preferentially ribonucleic acid.

4.5. Crystal structure of ChYtm1-ChErb1Ct dimer

4.5.1. Crystallization trials of ChErb1-ChYtm1 show that binding of ChYtm1 does not prevent degradation of ChErb1

Crystals of Erb1$_{416-807}$ from yeast indicated that when Nop7-Erb1 dimer was used for crystallization experiments Erb1 underwent sever degradation in a segment between Nop7 binding site and WD40 domain. Because that area of Erb1 (residues 383-419) was described to bind Ytm1, we expected that stable association of Ytm1 (or ChYtm1 in case of *C. termophilum* complex) would occlude the sequence that was recognized by the protease(s) present in the crystallization drop. Upon successful reconstitution of ChErb1-ChYtm1 dimer we attempted to produce crystals of the complex. Initial screenings yielded rod-shaped, flexible crystals that diffracted poorly up to 3.2 Å (Fig. 4.27a and Table 4.4).

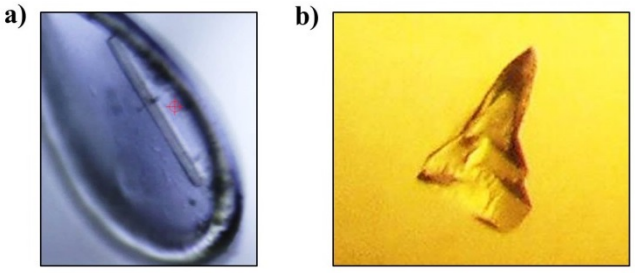


Figure 4.27. Crystallization of ChErb1-ChYtm1.
a) Crystallization screenings of ChErb1-ChYtm1 yielded rod-shaped crystals that grew in 15% PEG 4000, 0.1M sodium citrate pH 5.6, 0.2M Ammonium Sulfate and diffracted up to 3.2Å.
b) Crystallization trials of ChNop7-ChErb1-ChYtm1 trimer resulted in crystals 0% PEG 8000 and 0.1M Hepes pH 7.5 and diffracted at 2.1Å.

Diffraction pattern indicated primitive hexagonal crystals (space group P 6_5 2 2). Thanks to high symmetry of the crystal and in spite of the poor diffraction we were able to process the data and calculate the unit cell parameters.

	ChErb1$_{435-801}$-ChYtm1	**ChErb1$_{432-801}$-ChYtm1**
Wavelength (Å)	0.979	0.976
Resolution range (Å)[a]	48.82 - 3.1 (3.31 - 3.1)	48.53 - 2.1 (2.175 - 2.1)
Space group	P 65 2 2	P 21 21 2
Unit cell		
a, b, c (Å)	169.11, 169.11, 154.04	86.443, 108.149, 108.618
α, β, γ (°)	90, 90, 120	90, 90, 90
Total reflections	944854 (173208)	390408 (38570)
Unique reflections	24180 (4313)	60041 (5923)
Multiplicity	39.1 (40.2)	6.5 (6.5)
Completeness (%)	100.00 (100.00)	99.90 (99.97)
Mean I/sigma(I)[b]	17 (2.5)	18.69 (2.18)
Wilson B-factor	96.98	44.67
R-merge	0.282 (2.838)	0.05838 (0.924)
R-meas[c]	0.285	0.06359
CC1/2	0.998 (0.781)	0.999 (0.757)

Table 4.4 Data collection statistics of ChErb1Ct-ChYtm1.
[a] Statistics for the highest-resolution shell are shown in parentheses.
[b] Mean [$I/s(I)$] is the average of the relation between the intensity of the diffraction and the background.
[c] $R_{meas} = \{\Sigma_{hkl} [N/(N-1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle|\} / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and N is the multiplicity of reflection *hkl*.

Initial phasing using β-propeller models showed that two WD40 domains were present in the asymmetric unit (Table 4.5). Surprisingly, upon model building and manual inspection, we observed that the crystals contained full-length ChYtm1 and only the C-terminal β-propeller domain of ChErb1 thus indicating that the binding of ChYtm1 would not prevent degradation of ChErb1 during crystallization trials. In P 6_5 2 2 space group we were able to trace residues 13-487 of ChYtm1 and 427-801 of ChErb1 (ChErb1$_{427-801}$) with exceptions of several regions likely to be flexible.

	ChErb1$_{435-801}$-ChYtm1	ChErb1$_{432-801}$-ChYtm1
R-work[a]	0.1936 (0.2940)	0.1639 (0.2386)
R-free[b]	0.2390 (0.3557)	0.2132 (0.2703)
Number of non-hydrogen atoms	6170	6631
macromolecules	6168	6279
ligands	2[c]	26[d]
water	0	326
Protein residues	800	817
RMSD[e] **(bonds)**	0.006	0.008
RMSD[e] **(angles)**	1.26	1.18
Ramachandran favored (%)	94	96
Ramachandran outliers (%)	0.13	0.25
Clashscore	11.87	6.77
Average B-factor (Å)	102.40	63.20
Macromolecules (Å)	102.40	63.40
Ligands (Å)	96.10[c]	76.70[d]
Solvent (Å)		57.80

Table 4.5 Refinement statistics ChErb1Ct-ChYtm1.
[a] R-work = S_{hkl} {[$F_{obs}(hkl)$] - [$F_{calc}(hkl)$]} / S_{hkl} [$F_{obs}(hkl)$], where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.
[b] R-free corresponds to R_{factor} calculated using 2 % of the total reflections selected randomly and excluded during refinement.
[c] Ligands: Chloride ion
[d] Ligands: glycerol, ethylene glycol, ethanol
[e] RMSD is the root mean square deviation

4.5.2. ChYtm1-ChErb1$_{432-801}$ dimer crystallizes in P 2_1 2_1 2 space group in presence of ChNop7

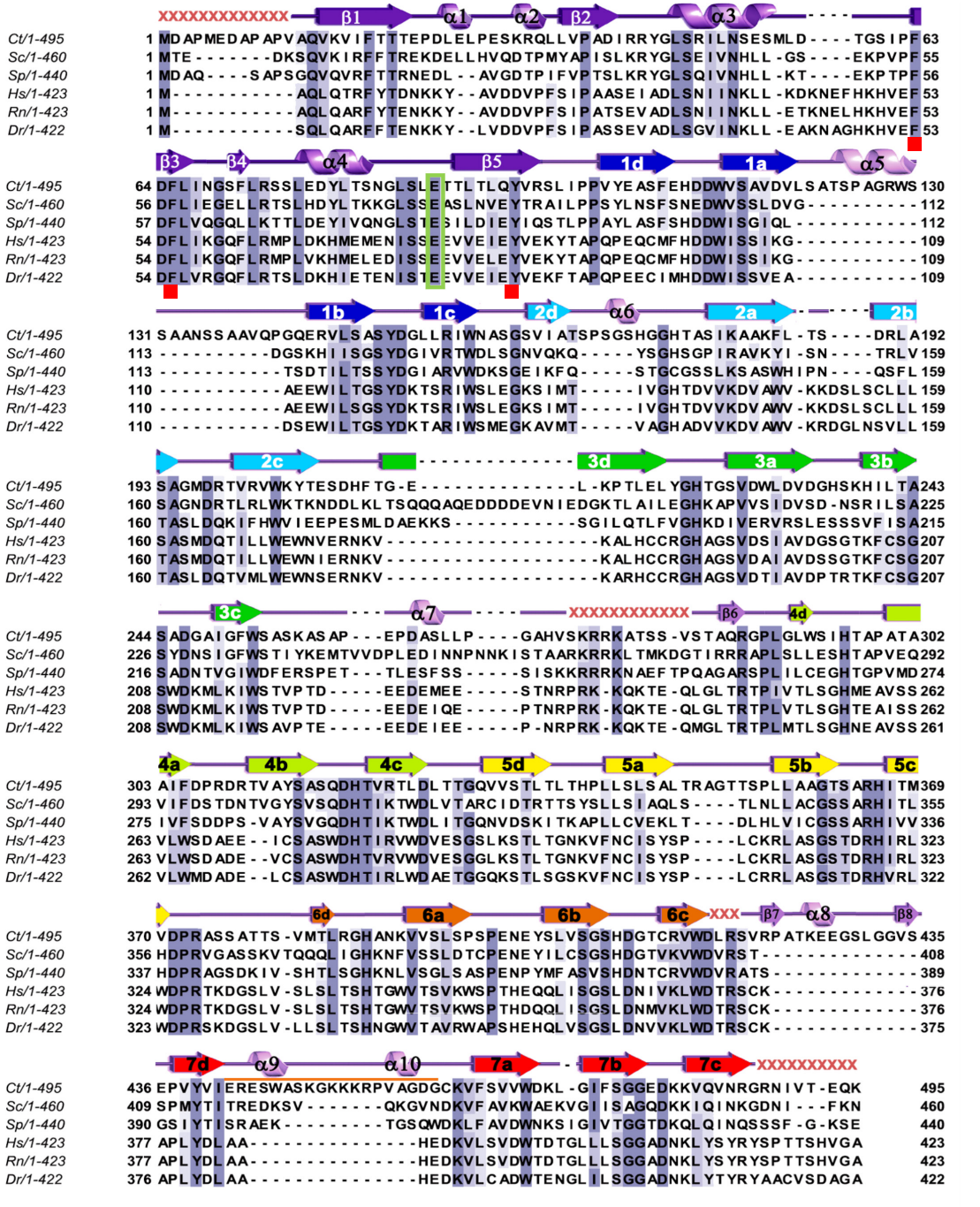
When expression and purification of ChNop7 were carried out with success, the complete heterotrimer was used in crystallization trials. After 3 to 5 days we observed crystal growth in several conditions and we were able to collect a dataset at 2.1 Å (Fig. 4.27b and Table 4.4). Although, the space group and unit cell parameters

differed from those grown in presence of ChYtm1-ChErb1 only, asymmetric unit content resulted to be identical. However, in this case, spatial organization of ChYtm1-ChErb1Ct dimer was different and crystal packing did not involve the most amino-terminal region of the β-propeller of ChErb1. Hence, we were able to build model for residues 12-487 of ChYtm1 and 432-801 of ChErb1 ($ChErb1_{432-801}$). Proteolytic cleavage in a site preceding β-propeller domain of ChErb1 was not prevented neither by binding of ChNop7 nor ChYtm1. Moreover, when PMSF 0.5mM was added to the crystallization drop we did not observe any crystal growth meaning that random proteolysis was necessary for the nucleation to occur. Since P 2_1 2_1 2 crystals yielded the best diffraction pattern and diffracted at highest resolution, we used the resulting model for all the structural analyses described in following chapters (Table 4.5).

4.5.3. Structural details of ChYtm1

Because no previous information regarding architecture of Ytm1 was available we analyzed the structure of the protein. ChYtm1 folded into two domains, a small N-terminal Ubiquitin-like Domain (UBL) followed by a seven-bladed β-propeller. Sequence alignment showed that Ytm1 from *C. thermophilum* presented insertion that were traced in the electron density map (Fig. 4.28).

Figure 4.28. Multiple sequence alignment of Ytm1.
Sequences corresponding to Ytm1/Wdr12 member from *Chaetomium thermophilum* (Ct),*Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp), *Homo sapiens* (Hs), *Rattus norvegicus* (Rn) and *Danio rerio* (Dr) were aligned. Secondary structure elements as seen in the final model of ChYtm1 are represented above the alignment. β-strands are marked with arrows, α-helices are depicted with coils.. Conserved residues are marked with shadows. The part of the protein corresponding to UBL domain is depicted in purple. Sequential WD repeats are colored and numbered by β-strands from 1d to 7c. Red squares show conserved residues within UBL that maintain the hydrophobic core of the fold. Conserved glutamic acid proposed to bind MIDAS domain of Rea1 is marked with a green box. The sequence corresponding to the knob-like structure is shown with orange line


Ct/1-495
Sc/1-460
Sp/1-440
Hs/1-423
Rn/1-423
Dr/1-422

4.5.3.1. Ubiquitin-like domain of ChYtm1

The amino-terminal sequence of Ytm1 protein has been annotated as NLE domain (Pfam: PF08154) and the crystal structure of ChYtm1 showed that the first 97 residues folded into an ubiquitin-like domain formed by two short α-helices and four antiparallel β-strands. In general the sequence of the UBL within Ytm1/Wdr12 family is poorly conserved and only residues that maintain the fold are invariable. The domain appears attached to the bottom side of the β-propeller (Fig. 4.29a). The inner face of the β-sheet contains several fully conserved residues (F63, F65 and Y95) (red squares in Fig. 4.28) that maintain a network of hydrophobic interactions with the first α-helix (residues 43-51) (Fig. 4.29b). Interestingly, the position of UBL with respect to the β-propeller domain was different between P 6_5 2 2 and P 2_1 2_1 2 space groups indicating conformational flexibility of the amino-terminal motif of ChYtm1, as discussed later.

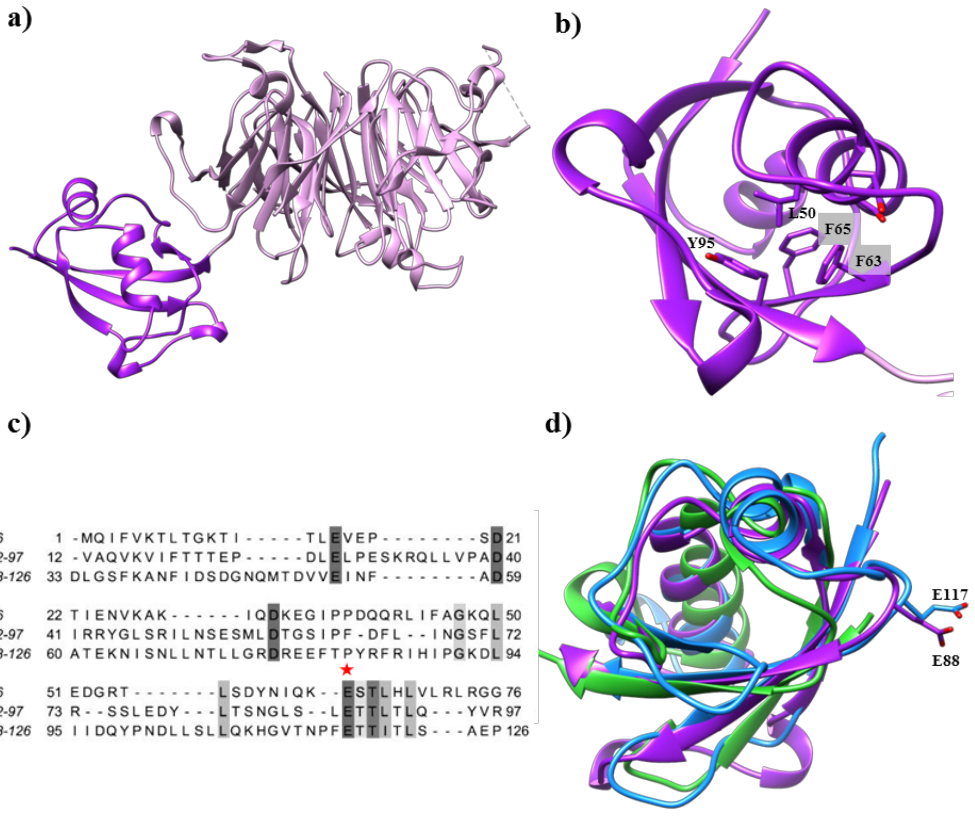


Figure 4.29. Ubiquitin-like domain of Ytm1.
a) Firs 97 residues of ChYtm1 fold into UBL domain (purple) that is attached to the bottom side of a seven-bladed β-propeller (pink). b) Conserved buried residues establish a network of hydrophobic interactions that maintain UBL fold. c) Sequence alignment of ubiquitin (residues 1-76), UBL domain of ChYtm1 (residues 12-97) and UBL of ChRsa4 (residues 33-126) shows that only few residues are conserved between the proteins. Glutamic acid important for function of ChYtm1 and CHRsa4 is marked with red asterisk. d) Superimposition of the structures of UBL of ChYtm1 (purple) with ubiquitin (PDB:UBQ1,green) and UBL of ChRsa4 (PDB:4WJS, blue) shows that the ubiquitin-like fold is well preserved. Side chains of glutamic acid involved in binding to MIDAS are shown for ChYtm1 and ChRsa4.

Although the structure of the domain highly resembles the fold of ubiquitin (RMSD between both structures is 1.2 Å), the sequence between both proteins is mostly dissimilar, thus proving that the ubiquitin-like motif is difficult to predict from the sequence (Fig 4.29c). At the same time UBL of ChYtm1 superimposes well with the N-terminus of Rsa4 from *C. thermophilum* (Fig. 4.29d) (RMSD from iPBA server: 1.75Å) although only few residues are fully conserved between both proteins, mainly those that maintain the hydrophobic core of UBL. Similarly to Rsa4, ChYtm1-Nt contains a conserved glutamic acid (E88) that mediates interaction with MIDAS domain of Rea1. Moreover, like in case of ChYtm1, two different orientations of UBL were seen in crystal structure of Rsa4. This flexibility and structural similarity between both UBLs point out toward a similar role of the N-terminal region from Ytm1 and Rsa4 in ribosome biogenesis as previously suggested by functional studies.

4.5.3.2. Ubiquitin-like domain shows conformational flexibility

As already mentioned, one of the main differences that we found between the two types of crystals of the dimer ChYtm1-ChErb1Ct was the relative position of UBL (Fig. 4.30a). To better understand this phenomenon we analyzed the interactions that are present between the N-terminal and β-propeller domains of ChYtm1 in each arrangement. Whereas the part of UBL (L66-F71) that contacts the extension from the blade 6 of the propeller (S4220-R422) does not change significantly from one space group to another, the loop that connects the first two

strands of UBL (T21-K32) appears much closer to the bottom face of WD40 domain in P 6_5 2 2 space group. This approximation pushes R419 toward N400 and induces a 90º rotation of the side chain of Y402 that in consequence coordinates with T22 from UBL (Fig. 4.30b).

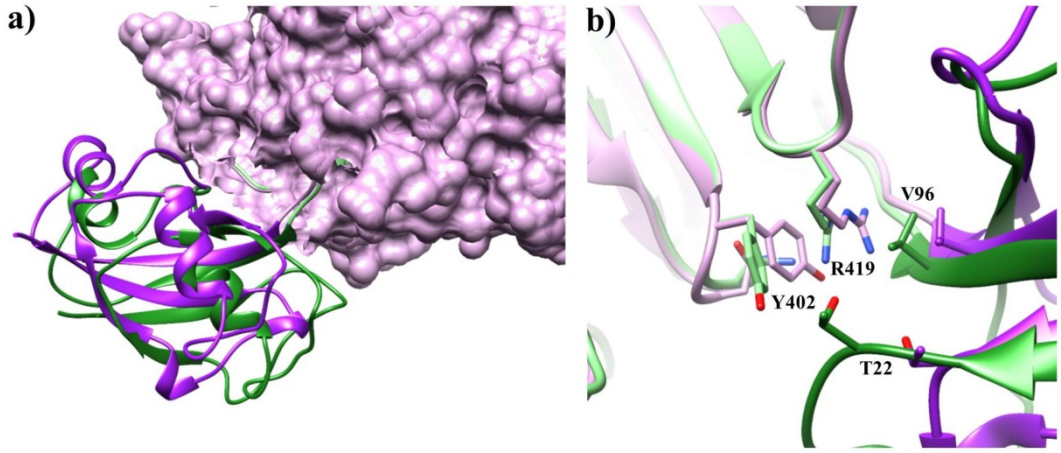


Figure 4.30. Ubiquitin-like domain of Ytm1 is flexible.
Model obtained with P 6_5 2 2 crystal is represented in green and the one corresponding to P 2_1 2_1 2 space group is shown in pink/purple. a) Relative position of UBL changes between different types of crystals. b) When the last β-strand of UBL gets closer to the propeller, in P6_5 2 2 crystals, V96 pushes R419 towards Y402 inducing rotation of its side chain so that it can establish contacts with T22 from UBL that orientates much closer to the propeller than in P 2_1 2_1 2 crystals.

4.5.3.3. β-propeller domain of ChYtm1

The carboxy-terminal domain of ChYtm1 is formed by 7 WD repeats that arrange in a tight propeller structure (Fig. 4.31a). Only WD6 contains a fully conserved WD di-peptide at the end of strand "c". The core of the domain highly resembles other 7-bladed β-propellers and superimposes well (RMSD ≤ 2Å) with more than 100 proteins from PDB as calculated by Dali Server (Fig. 4.31b). The main difference can be attributed to the loops between β-strands (blue segments in Fig. 4.31b) that, in several regions, fold into additional secondary structures.

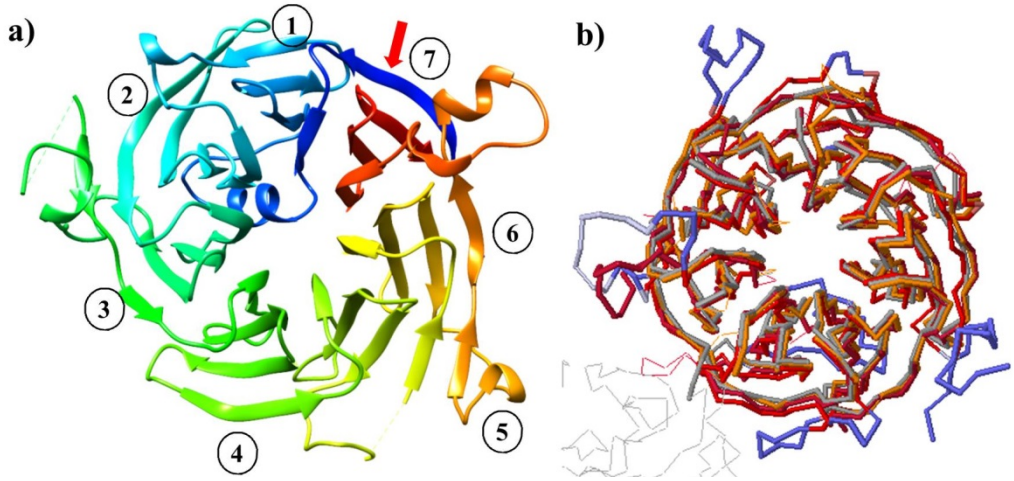


Figure 4.31. β-propeller domain of ChYtm1.
a) Top view of the β-propeller of ChYtm1 represented with rainbow coloring. Blade numbering is indicated. The velcro-like 1+3 closure of the domain is marked with red arrow. Blades 2 and 5 contain additional β-strand in their folds. A short α-helix inserted between blades 1 and 2 can be seen to partially occlude the central channel of the propeller. b) Top 3 most similar structures from PDB (as calculated by Dali Server) were superimposed to the one of ChYtm1$_{100\text{-}487}$. (Chains were colored as follows: 4d6v:A- dark red, 2ymu:A-grey, 4esg:A-orange)

Noticeably, some of these extended loops seem to be fungi specific and are lost during evolution (Fig. 4.32). In the bottom part of the domain, arising from the first blade appears a short α-helix (residues 125-131) that partially occludes the central aperture (A in Fig.4.32).

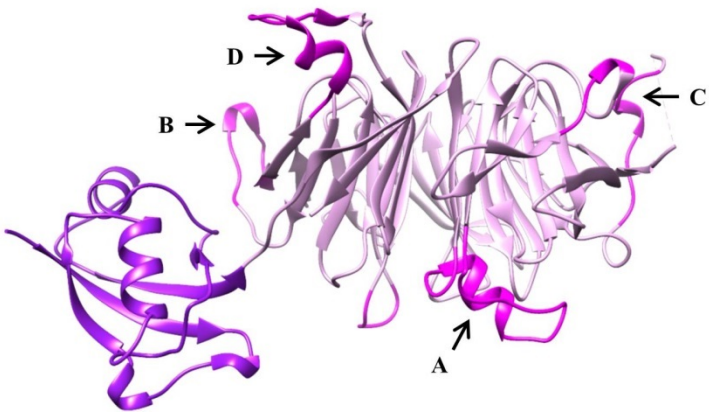


Figure 4.32. Insertions within β-propeller domain of ChYtm1.
C-terminal domain of ChYtm1 (light pink) contains several extensions that fold into additional motifs on the surface of the propeller (magenta) and tend to disappear during evolution. A: residues 121-144, B: 422-436, C: 260-269, D: 444-461.).

The segment can only be found in certain group of fungi and is absent in yeast or higher Eukaryotes. A long non-conserved loop between WD6 and WD7 (418-436) contains a β-strand that becomes the fifth strand of blade 5 (B in Fig. 4.32). Only the initial fragment of this "c-d" loop can be observed in yeast and mammals and, as seen in the structure, establishes important contacts with the UBL thus modulating its position. The second blade is also built by five and not four β-strands, but in this case the strand "e" is conserved in higher Eukaryotes. In Fungi it is covered by additional unstructured fragment that covers the area between blades 2 and 3 (C in Fig. 4.32). The last extended loop connects blades 6 and 7 (residues 444-461) and forms a protrusion on the upper face of the propeller (D in Fig. 4.32). This knob-like structure has been previously observed in Asc1 protein which is known to bind 40S.

When we analyzed the conserved patches on the surface of the propeller it was clearly noticeable that the least variable area corresponded to the ChErb1-interacting face (Fig. 4.33). Regarding the most conserved feature of the β-propellers, the aspartic acid residue from the "b-c" loops (b-c Asp), in case of ChYtm1 is present in 6 blades. Only the fifth blade lacks the aspartic acid and possesses a conserved arginine that, interestingly, is located in the area that is close to the β-propeller of ChErb1 although the residue itself does not seem to be directly involved in the interaction.

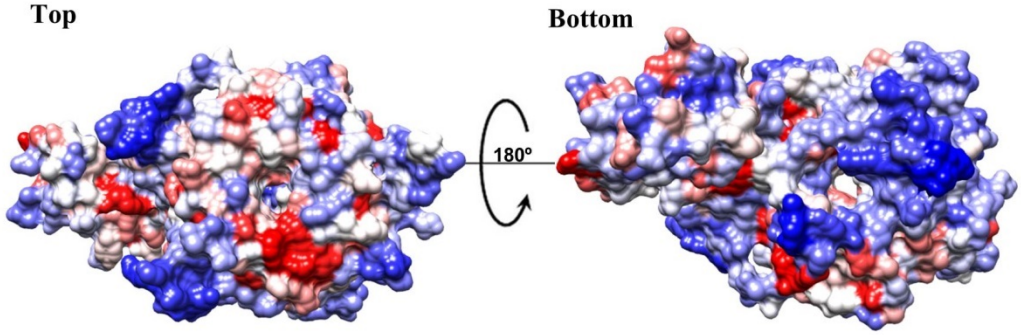


Figure 4.33. Conserved residues on the surface of ChYtm1.
Surface view of ChYtm1 rendered by conservation as calculated from the multi-alignment made with ClustalO. Top face of the propeller contains more fully conserved residues than its bottom surface

The structural triad, essential for the β-propeller folding and described previously, is present in all, except one, blades that have conserved "b-c" Asp. The WD repeat 7 has a long insertion between strands "d" and "a" which lacks the GH motif in the "d-a" loop. Instead, it has an arginine (R443) that mimics the role of the histidine in the triad formation and makes electrostatic contacts with S475 and D479 that allow proper blade organization. Curiously, this arrangement is found in fungi but higher Eukaryotes recover the GH motif and the canonical histidine-containing triad.

Another peculiarity of the β-propeller of ChYtm1 is the unusual length of the β-strands "c" and "d" from the second blade (2c and 3d). These strands project toward the bottom part of the first blade. R145 from the beginning of 1b makes a salt bridge with E207 from 2c thus forming a clasp that completely closes the breach between blades 1 and 2.

4.5.4. β-propeller domains of Erb1 from *C. termophilum* and *S. cerevisiae* are highly similar

The C-terminal domain of ChErb1 superposed well with the previously solved β-propeller of Erb1 from yeast (Fig. 4.34a) (RMSD from iPBA server: 0.98Å). The most dissimilar area between both proteins corresponded to the insertion that appeared within blade two, as described previously for *S. cerevisiae,* and an adjacent loop that connected strands "2d" and "2a" (marked with red arrow in Fig. 4.34a). It is also worth noting that the most flexible fragment of the insertion that could not be traced due to the missing electron density was much shorter in case of *C. termophilum* (17 residues) when comparted to the same region in Erb1 from yeast (36 residues). As mentioned previously, the low-resolution structure from P 6_5 2 2 crystals allowed us to trace an additional fragment of the β-propeller that extended towards the N-terminal region of the protein. This probably flexible segment (residues 427-432) became rigid and traceable because it packed against the same chain from the symmetrically related monomer. Interestingly, the large positively charged area described for *S. cerevisiae* Erb1Ct is maintained in ChErb1Ct which had been shown to bind RNA in vitro (Fig. 4.34b).

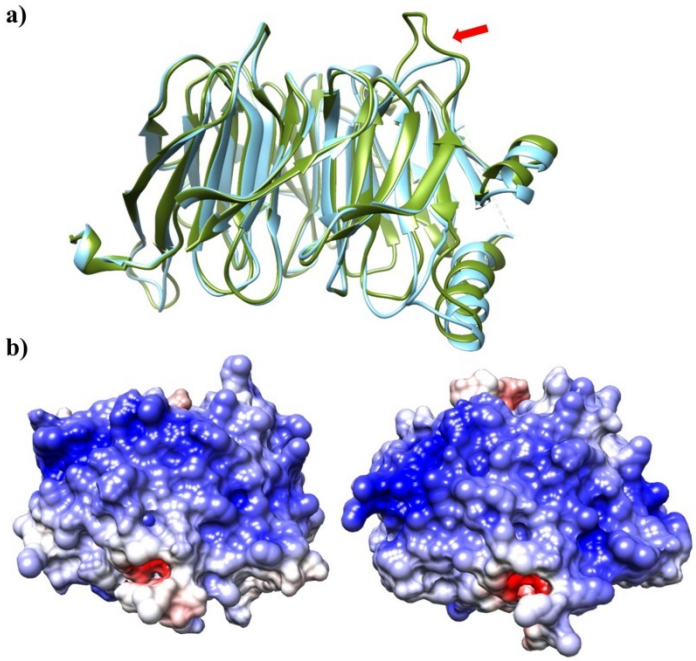


Figure 4.34. ChErb1$_{432-801}$ superposes well with its counterpart from yeast.
a) The structure of ChErb1$_{432-801}$ (blue) is well conserved when compared to Erb1$_{416-807}$ (green). Main difference between both domains resides in the "2d-2a" loop (red arrow) and the length of the insertion. b) The positively charged area on the surface of Erb1Ct is well conserved between species as shown for *C. thermophilum* (left) and *S. cerevisiae* (right).

4.5.5. ChYtm1 interacts with the β-propeller domain of ChErb1

The unexpected disposition of full-length ChYtm1 bound to the C-terminal domain of ChErb1 in the asymmetric unit led us to further investigate the possibility of a direct interaction between both β-propellers. Since previously published data suggested that the binding of Ytm1 to Erb1 in yeast did not involve the β-propeller of Erb1 (Tang et al., 2008), we first computationally checked whether the dimer was stable in solution.

Assembly Summary			
Multimeric state	2	**Surface area (Å^2)**	32738.2
Copies in unit cell	4	**Buried area (Å^2)**	2457.6
ΔG^{int} (kcal/mol)[a]	-8.5	**$T\Delta S^{diss}$ (kcal/mol) c**	13.8
ΔG^{diss} (kcal/mol)[b]	4.3	**Symmetry number**	1
Formula	AB		
Composition	AB		
Dissociation pattern	A + B		

Table 4.6. The most probable assembly as calculated by PISA Server.
Stability of the dimer in solution was computationally confirmed.
[a] ΔG^{int} indicates the solvation free energy gain upon formation of the assembly
[b] ΔG^{diss} indicates the free energy of assembly dissociation
[c] $T\Delta S^{diss}$ indicates the rigid-body entropy change at dissociation

PISA server confirmed that the complex formation was energetically favorable, being the buried area between both monomers 2458Å^2 (Krissinel & Henrick, 2007). According to PISA 43 residues of ChYtm1 and 34 from ChErb1 were engaged in the dimer formation (Table 4.6).

4.5.5.1. High-affinity binding occurs between ChYtm1 and ChErb1$_{432-801}$ in vitro

We performed *in vitro* binding assays using ChYtm1 and truncated ChErb1 (residues 432-801). Both proteins were mixed and injected into size exclusion chromatography column. ChYtm1 and ChErb1$_{432-801}$ co-eluted from gel filtration and we could observe approximately 1:1 binding on SDS-PAGE (Fig. 4.35a). Elution volume indicated that both proteins formed a dimer. Moreover, we performed a pull down assay using Glutathione Sepharose beads. First, GST-tagged ChYtm1 or GST alone were attached to the resin and then 6xHis-tagged ChErb1$_{432-801}$ was added to the beads. After extensive washes, the proteins were eluted with glutathione-containing buffer. Whereas, we did not observe any 6xHis-ChErb1$_{432-801}$ in the control elution, it was present in the sample that contained GST-ChYtm1 clearly suggesting that the β-propeller of ChErb1 was stably associating with ChYtm1 (Fig. 4.35b).

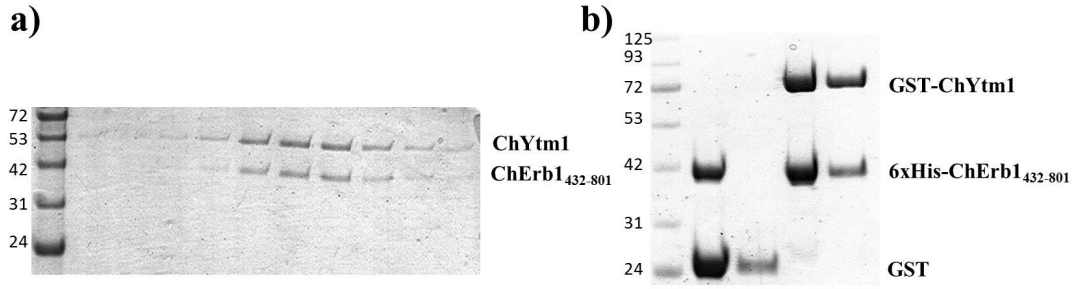
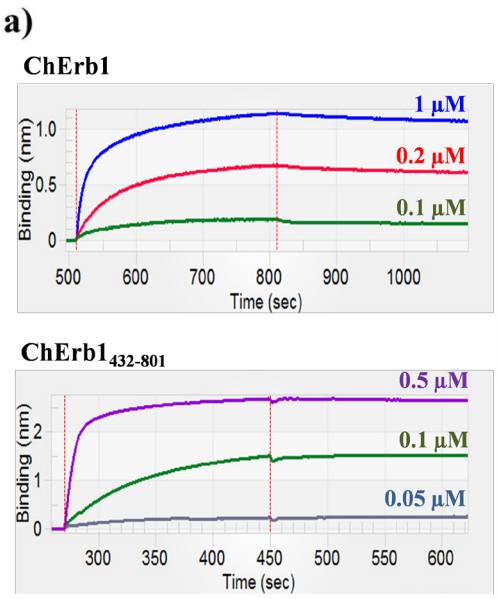
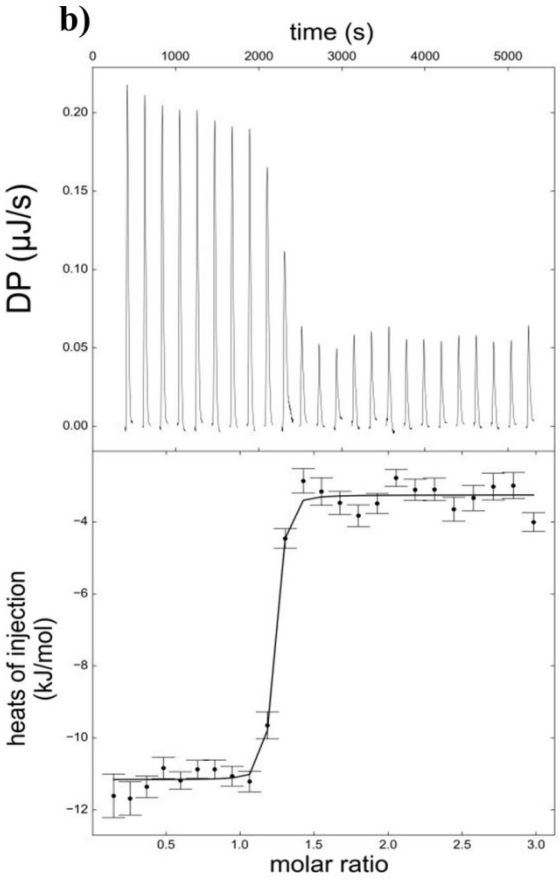


Figure 4.35. β-propeller of ChErb1 and ChYtm1 interact in vitro.
a) SDS-PAGE of the fractions corresponding to gel filtration elution profile where one wide peak was observed. Both, ChYtm1 and ChErb1$_{432-801}$ were found to co-elute together. b) Pull-down experiment showed that 6xHis-tagged ChErb1$_{432-801}$ co-purified with GST-tagged ChYtm1 but not with GST alone on Glutathione Sepharose beads. 1: Input, 2: Elution

Because ChYtm1 and ChErb1$_{432-801}$ held together in gel filtration we expected the binding to occur with a good affinity. In order to quantify the strength of the interaction we calculated the dissociation constant (K_D) using biolayer interferometry. GST-ChYtm1 was immobilized on Anti-GST biosensor and then increasing concentrations of either ChErb1 or ChErb1$_{432-801}$ were used in association and dissociation steps. We observed a good binding between ChYtm1 and both ChErb1 constructs and we could measure the affinity. Comparable K_D of ~4 nM and ~3 nM, for ChErb1 and ChErb1$_{432-801}$ respectively, indicated that the binding occurred with high affinity in both cases (Fig. 4.36a). Moreover, since the β-propeller alone bound ChYtm1 with the same strength as the full-length protein, the N-terminal portion of ChErb1 was dispensable for the stable association to occur.





K_a = 1.088e+08 M^{-1}	**ΔG** = -10.780 kcal/Mol
K_D = 9.193e^{-09} M	**TΔS** = 8.550 kcal/Mol
ΔH = -2.230 kcal/Mol	**ΔS** = 29.165 cal/Mol*K

Figure 4.36. β-propeller of ChErb1 and ChYtm1 interact in vitro.
a) Biolayer Interferometry experiment showed that both ChErb1 (top) full-length and ChErb1$_{432-801}$ (bottom) bound ChYtm1 with a good affinity. b) ITC was done to validate high-affinity binding between ChYtm1 and ChErb1$_{432-801}$. Raw heat of each injection is shown in the upper panel. Curve fitting allowed to confirm that the binding affinity was in low nanomolar range (K_D ~10nM). Calculated thermodynamic parameters for the ITC experiment are shown in the table.

Next, we performed an isothermal titration calorimetry (ITC) assay to get an insight into the kinetics of the interaction between ChYtm1 and ChErb1$_{432-801}$. The association of the proteins resulted in heat release and the measured peaks were integrated and thermodynamic parameters calculated. Strong interaction, with K_D in a low nanomolar range (~9nM), was confirmed to occur between ChYtm1 and the carboxy-terminus of ChErb1 (Fig. 4.36b).

4.5.5.2. Details of ChYtm1-ChErb1$_{432-801}$ dimer interface

As calculated by PISA Server, there are 18 hydrogen bonds and 11 salt bridges across the dimer interface. Manual inspection of the model confirmed that the dimer was mainly maintained by electrostatic forces although several hydrophobic regions were also involved in the interaction. ChYtm1 uses the extensive top face of its β-propeller to establish contacts with the bottom part and the circumference of ChErb1Ct (Fig. 4.37).

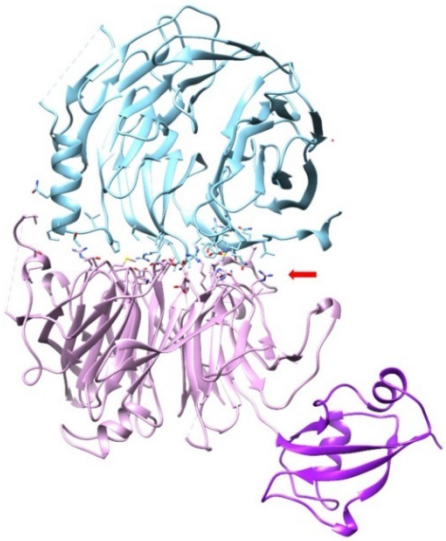


Figure 4.37. Top face of the β-propeller of ChYtm1 binds ChErb1.
Ribbon representation of the dimer shows that ChYtm1 (pink) binds to the C-terminal domain of ChErb1 (blue) through the top surface of the propeller (red arrow). UBL domain of ChYtm1 (purple) does not participate in the interaction. Side chains of interacting residues are shown.

The manner in which both propellers interact is quite unusual as the axis of one of the domains is tilted 55º with respect to the other (Fig. 4.38).

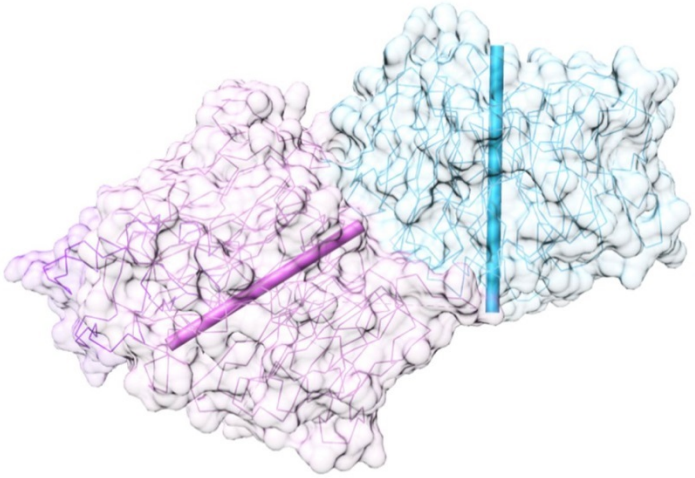


Figure 4.38. Particular arrangement of WD40 domains within the dimer. The bottom edge of ChErb1Ct inserts into the central cavity on the tope face of ChYtm1. The central axes of the β-propellers (shown as bars) form an angle of 55º.

First interacting region includes the last β-strand ("1d") of blade 7 in ChErb1Ct that contacts loop "6d-6a" and a long extension that appears between strands "7d" and "7a" of ChYtm1 (the knob formed by residues 444-460) (A in Fig.4.39). Since this extended loop is not conserved in other Fungi or higher Eukaryotes it is possible that the way both proteins interact at this point is *Chaetomium* specific. Blade 7 of ChErb1 is also engaged in another contact through its "7a-7b" loop, which contains E785 that forms salt bridge with fully conserved H320 from ChYtm1.

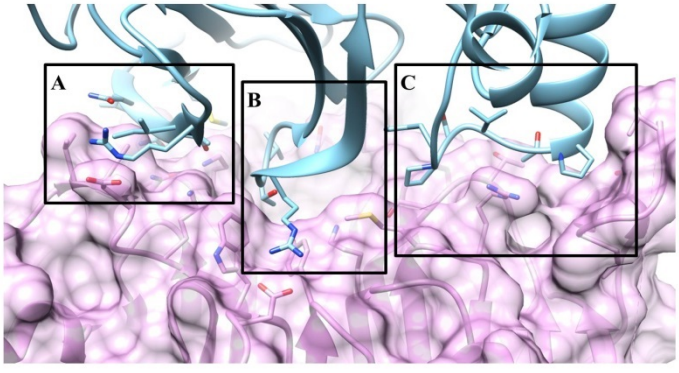


Figure 4.39. Three areas of ChErb1$_{432\text{-}801}$ contact the β-propeller of ChYtm1. A) Strand 1d from the blade 7 contacts the knob that appears between blades 6 and 7 of ChYtm1. B) Conserved loop "c-d" from blade 1 of ChErb1 binds to the central channel of ChYtm1. C) The insertion from blade 2 of ChErb1 mediates binding to an extended loop from ChYtm1 on one side of the propeller.

The entrance of the central tunnel of ChYtm1 serves as place of docking for a loop between strands "1c-2d" from the first blade of ChErb1Ct (481-486) (B in Fig. 4.39). The loop contains three well conserved residues: E481, T484 and R486 that establish a network of electrostatic interactions with also conserved amino acids from blades 1, 2, 3 and 7 of ChYtm1. A salt bridge is formed between R486 of ChErb1 and D112 from ChYtm1. Two additional residues of ChYtm1: W113 and Y151 ensure the proper orientation of R486 side chain thus promoting electrostatic contacts. Interaction with ChYtm1 actually induces conformational reorganization that allows salt bridge formation because in the structure of Erb1Ct alone (from *S. cerevisiae*) the residue corresponding to R486 (R470 in S. cerevisiae) is coordinated by E481 (E465 in *S. cerevisiae*) (Fig. 4.40). Upon binding, E481 forms hydrogen bonds with K181 and Y151 from ChYtm1 and the side chain of R486 is forced to move toward D112.

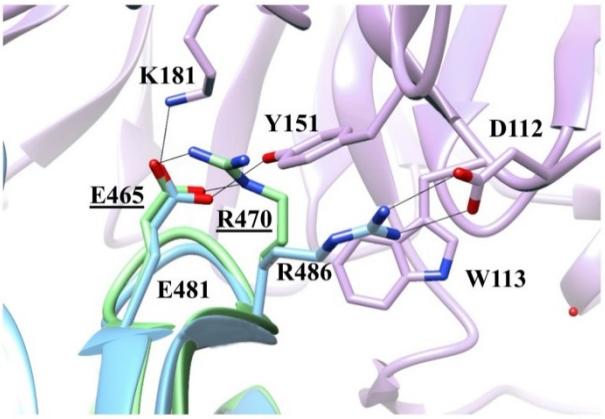


Figure 4.40 Residues involved in R486-D112 salt bridge formation.
ChYtm1 is shown in pink and ChErb1 in blue. Superimposition of monomeric β-propeller of Erb1 from *S. cerevisiae* (in green) shows the conformation switch that forces R486 (R470 in yeast) toward salt-bridge formation. Electrostatic interactions are shown as black lines. Labels corresponding to residues from yeast Erb1 are underlined.

At last, our crystal structure showed the importance of the insertion that appeared within the second blade of the β-propeller of Erb1. We observed that in ChYtm1-ChErb1Ct dimer this region provided additional surface of interaction because it made important contacts with the loops from blade 2 and 3 of ChYtm1 and with another long extension, between strands "3c" and "4d" that projects out of

the plane formed by the top surface of ChYtm1 thus increasing dimerization area (C in Fig.4.39). In conclusion, the central part of the propeller of ChYtm1 provides a large docking surface for the bottom face of blades 1, 2 and 7 from ChErb1Ct that is additionally held in place by two lateral extensions from ChYtm1.

4.5.5.3. Electrostatic repulsion induced by R486E mutation affects the binding affinity

In order to validate the structural information provided by our model we attempted to prove that the β-propeller of ChErb1 played crucial role in binding to ChYtm1. We analyzed the conservation of the interfaces that participate in the dimer formation and we found an area that was fully conserved in both proteins (Fig. 4.41).

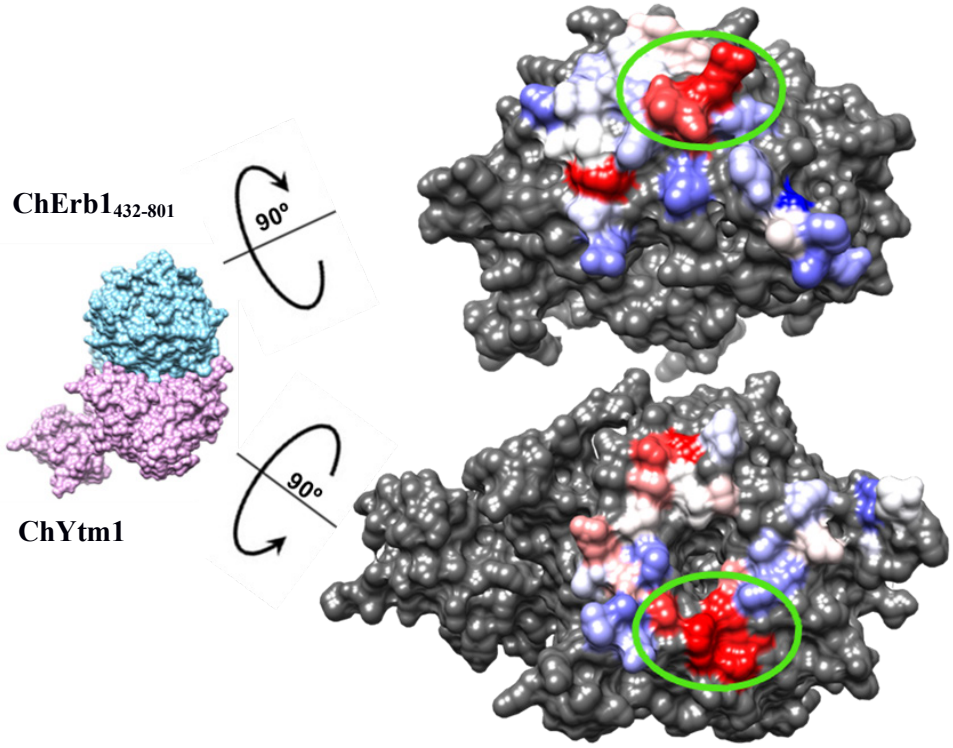
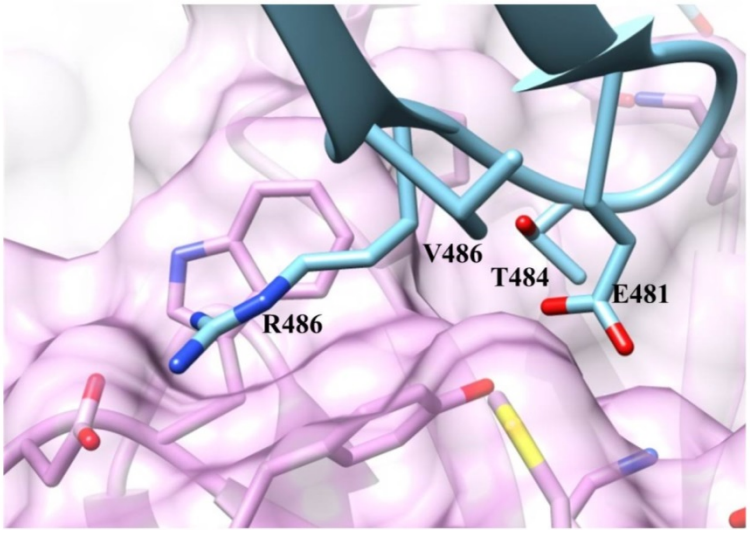


Figure 4.41. Conservation of the surface involved in ChErb1$_{432-801}$-ChYtm1 interaction.
The colored surfaces of ChErb1Ct (top) and ChYtm1 (bottom) show the area that mediates the binding between both molecules. The most variable residues are shown in blue and the most conserved one are depicted in red. Green circles mark the region that was chosen for mutant generation.

This region corresponded to the previously described salt bridge between R486 and D112 and adjacent residues. We designed a set of ChErb1 point mutations that altered the most conserved residues from that area of interaction (Fig. 4.42). Two of the mutants (E481R and T481E) expressed as inclusion bodies indicating misfolding of the nascent propeller. Those that were soluble upon expression were purified and assayed for their binding to ChYtm1. Initially, we checked whether any mutant of ChErb1 would lose its ability to co-elute with ChYtm1 in gel filtration thus indicating lack of binding or at least an important decrease in its affinity. Only one of the tested proteins, ChErb1[R486E], did not hold together with ChYtm1 in size exclusion column.



Mutant	S.	SE	K_D (M)
E481R	NO	X	-
E481D	YES	+	$5x10^{-9}$
T484E	NO	X	-
T484Q	YES	+	$3x10^{-9}$
R486A	YES	+	$7x10^{-9}$
R486E	YES	-	$249x10^{-9}$
V488W	YES	+	$7x10^{-9}$

Figure 4.42. "1c-2d" loop of ChErb1$_{432-801}$ was mutated to disrupt the binding.
Side chains of modified residues of ChErb1 (blue) are labeled and corresponding mutations are listed on the right. Soluble variants (YES in S. column) were injected with ChYtm1 in gel filtration (SE column: X: not checked, +: bound, -: no binding) and interaction affinity was measured by interferometry.

Next, we carried out a pull-down experiment to assess if this point mutation was disrupting the binding completely and we observed that there was still a detectable amount of 6xHis-ChErb1[R486E] co-purifying with GST-ChYtm1 on Sepharose Glutathione beads suggesting that both proteins were still associating but not as stably as in the case of wild-type ChErb1. Finally we decided to quantify the decrease of affinity produced by the mutant (Table in Fig. 4.42). The K_D measured

by interferometry using ChYtm1 with full-length ChErb1[R486E] or ChErb1$_{432-801}$[R486E] resulted to be two orders of magnitude lower (K_D ~250nM) than ChErb1-ChYtm1 interaction described previously. At the same time we confirmed that other mutations did not significantly affect the binding and presented K_D values within wild-type range (~3-7 nM). Interestingly, disruption of the salt bridge by R486A mutation was not sufficient to weaken the binding (Erb1-R486A co-eluted with ChYtm1 in gel filtration) and only an electrostatic repulsion induced by R486E had an important effect on the interaction.

4.5.5.4. R468E mutation does not affect the β-propeller architecture

To further investigate the impact of R486E mutation on the binding we checked if the decreased affinity was not a side effect of aberrant β-propeller folding or its reduced stability in vitro but was directly related to the repulsion between E486 and D112 from ChYtm1. Using Thermofluor we confirmed that thermal stability of R486E β-propeller was not compromised and its melting temperature was comparable to the one of wild-type ChErb1$_{432-801}$ (Fig. 4.43a). We performed crystallization trials with ChErb1$_{432-801}$[R486E]-ChYtm1. Four very similar conditions yielded crystals of comparable morphology that was pH-dependent (Fig. 4.43b). The highest resolution was achieved for the crystals that appeared in the condition that contained 0.1 M Hepes pH 7.5 and 2M Ammonium Sulfate. A complete dataset allowed us to solve a 3.0Å structure of the dimer (Table 4.7) that verified that the overall architecture of the C-terminal domain of Erb1 was not altered in the context of the mutation (RMSD between ChErb1$_{427-801}$ and ChErb1$_{432-801}$[R486E] was 0.48Å). As expected, mutant protein still interacted with ChYtm1 *in vitro* at the concentrations used in the crystallization drops and the electron density for E486 from ChErb1 clearly showed a lack of coordination with D112 from ChYtm1 (Fig. 4.43c).

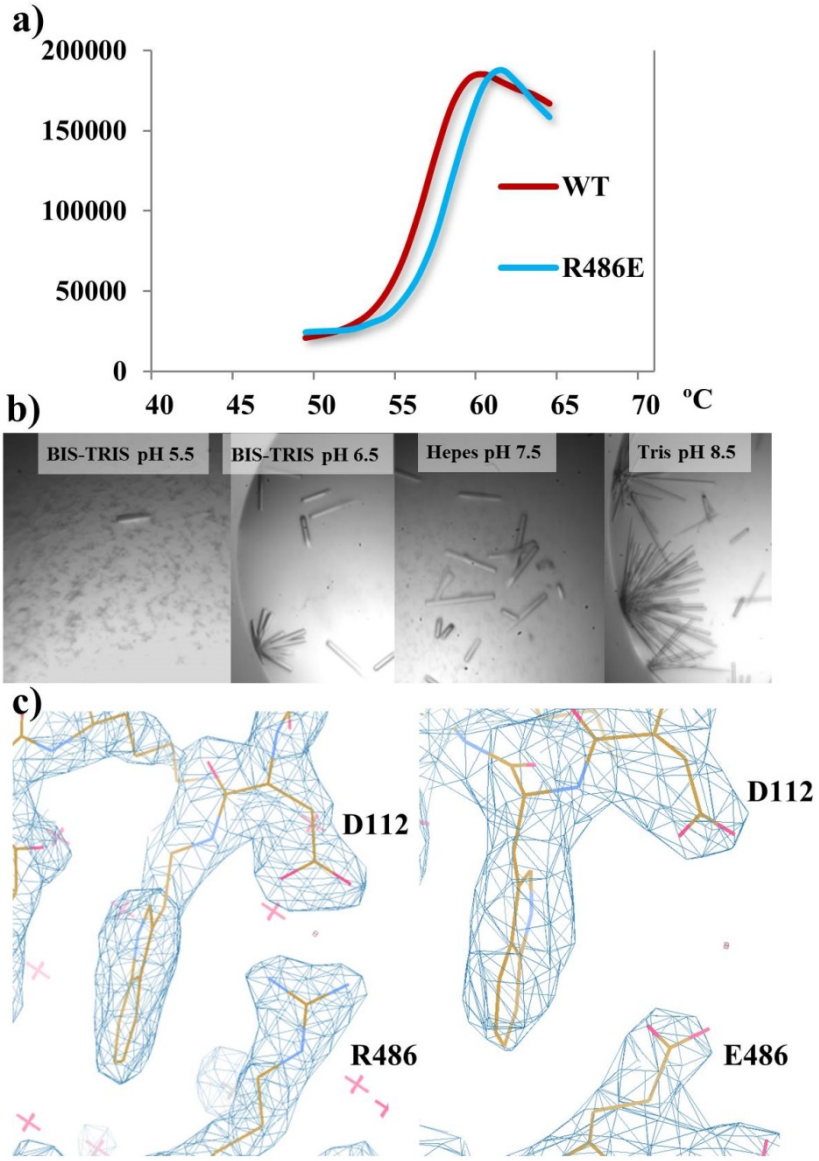


Figure 4.43. The stability and structure of ChErb1$_{432-801}$ are preserved in the context of R486E mutation.

a) Thermal stability of ChErb1$_{432-801}$[R486E] (blue) is equivalent to the wild-type protein (red). b) Crystals containing mutant ChErb1 and ChYtm1 grew in 2M ammonium sulfate supplemented with buffers at different pH c) Upon successful crystallization of ChErb1$_{432-801}$[R486E] we proved that the β-propeller structure was intact and the only difference between both structures was the lack of D112-R486 salt bridge (left) between D112 and E486 (right). Electron density is shown with 1.5σ.

Interestingly, ChYtm1-ChErb1$_{432-801}$[R486E] formed crystals with exactly the same symmetry and unit cell parameters as those of ChYtm1-ChErb1 in P 6_5 2 2 space group. It was surprising because, as mentioned in the section 4.5.4, the most amino-terminal end of the β-propeller of ChErb1 (residues 427-431) seemed to be involved in crystal contacts with the same region of the symmetrically related molecule. However, the sample used for the crystallization of the mutant-containing dimer did not have full-length ChErb1[R486E] but only the β-propeller itself (residues 432-801), thus lacked the N-terminal extension. Obviously there was no electron density in the area corresponding to 427-432 segment of ChErb1 which indicated that the loss of one of the crystal contacts did not impede crystal formation in P 6_5 2 2.

	ChErb1[R486E]$_{435-801}$-ChYtm1
Wavelength (Å)	0.979
Resolution range (Å)[a]	85.38-3.0 (3.18-3.0)
Space group	P 65 2 2
Unit cell	
a, b, c (Å)	170.75, 170.75, 155.66
α, β, γ (°)	90, 90, 120
Total reflections	332895 (55186)
Unique reflections	27367 (4333)
Multiplicity	12.2 (12.7)
Completeness (%)	100.00 (100.00)
Mean I/sigma(I)[b]	13.5 (1.6)
Wilson B-factor	87.72
R-merge	0.176 (1.862)
R-meas[c]	0.184 (1.940)
CC1/2	0.995 (0.606)
R-work[d]	0.2016 (0.3153)
R-free[e]	0.2621 (0.3914)
Number of non-hydrogen atoms	6182
macromolecules	6177
ligands[f]	5
water	0
Protein residues	804
RMSD[g] **(bonds)**	0.006
RMSD[g] **(angles)**	1.35
Ramachandran favored (%)	94
Ramachandran outliers (%)	0.13
Clashscore	13.72
Average B-factor (Å)	89.70
Macromolecules (Å)	89.70
Ligands[f] **(Å)**	84.70
Solvent (Å)	X

Table 4.7 Data collection and refinement statistics of ChErb1$_{432-801}$[R486E]-ChYtm1.
[a] Statistics for the highest-resolution shell are shown in parentheses.
[b, c, d, e, g] Like in Tables 4.4 and 4.5
[f] Ligands: chloride ion

4.6. Functional studies corroborate the structural findings.

By disrupting of the fully conserved salt bridge between R486 of ChErb1 and D112 of ChYtm1 we managed to weaken the binding between both proteins *in vitro*. We decided to check whether a similar mutation would have an effect in *Saccharomyces cerevisiae*, the model organism for studying eukaryotic ribosome biogenesis. We inserted erb1 gene into YCplac33 vector which we then used to complement null erb1 strain. Next, we cloned erb1 with its native promoter into YCplac111 and used it as a template to generate erb1R470E mutant which correspond to Erb1R486E from *C. termophilum*. Finally we investigated the impact of the mutation on the phenotype of yeast.

4.6.1. Erb1[R470E] impairs growth in yeast

We used Erb1 or Erb1[R470E] cloned into YCplac111 to investigate whether we could rescue the lethal phenotype of the null erb1 strain. For that reason we transformed Δerb1/YCplac33-Erb1 with either wild-type or mutant copy of Erb1-YCplac111. Once transferred to FOA-containing medium the cells lost YCplac33-Erb1 and their growth was dependent on the gene expressed from YCplac111 constructs. As shown by spot assay, at 30ºC, normal growth was fully restored by the wild-type Erb1 whereas Erb1[R470E] resulted in only slightly smaller colonies indicating a very mild growth defect (Fig 4.44).

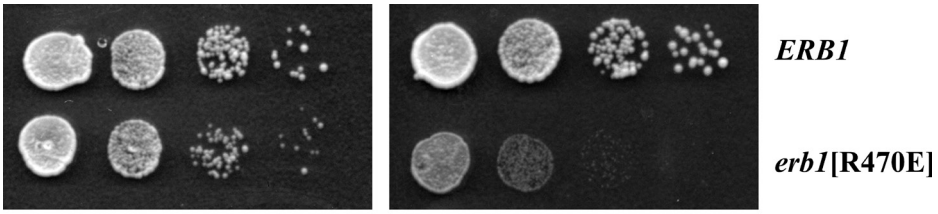


Figure 4.44. Erb1[R470E] impairs growth in yeast.
At 30ºC Δerb1 strain complemented with erb1[R470E] grew only slightly slower than the one carrying wild-type erb1, however a significant growth alteration in the context of the mutation was observed at 37ºC.

At 37ºC the strain complemented with Erb1 still grew normally but the one containing Erb1[R470E] clearly showed altered growth rate.

4.6.2. 60S subunit biogenesis is affected in the context of Erb1[R470E]

To further explore the impact of the mutation on the yeast cells we asked if the slower growth could be correlated to abnormalities during ribosome synthesis. We extracted the polysomes and subjected them to fractionation on sucrose gradients. As seen in the profiles, the wild-type Erb1 showed unaffected peaks corresponding to the ribosomal subunits, 80S assembly and translating ribosomes. On the other hand, there was a clear effect of Erb1[R470E] (Fig. 4.45). Once again, it resulted to be less prominent at 30ºC and became more sever at higher temperature. In both cases there is a visible decrease in 60S subunits and half-meres arising in the peaks corresponding to the 80S and in the polysomes. The profile of Erb1[R470E] at 37ºC shows much more drastic lack of functional ribosomes which seriously hinders translation.

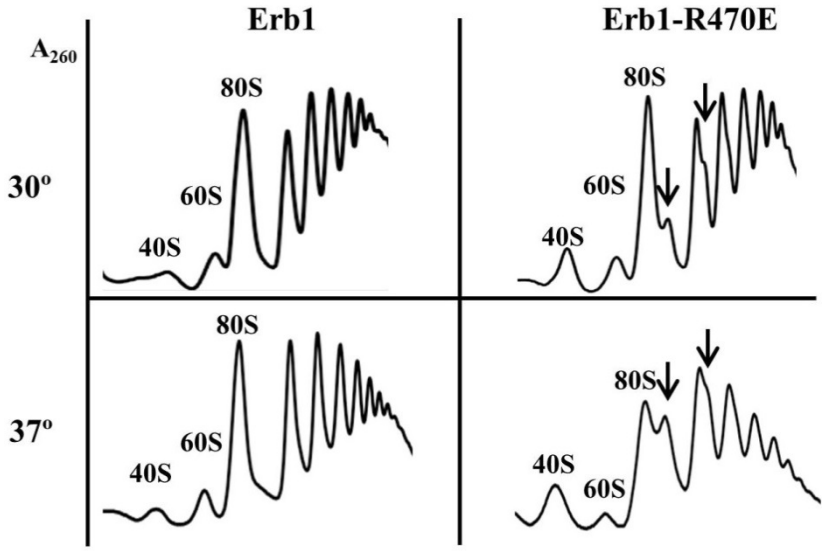


Figure 4.45. Erb1[R470E] results in altered 60S biogenesis.
At 30ºC the effect of the mutation was less prominent than at 37ºC. In both cases, there is a clear deficit of 60S subunit; halfmeres (black arrows) appear in 80S and polysomal fraction. Decreasing polysomes levels in Erb1[R470E] at 37ºC indicate altered translation.

To better quantify the reduction of LSU particles in the context of the mutation we analyzed the amount of each subunit in yeast grown at 37ºC and we observed that whereas in Erb1 the ratio 60S/40S was 2.2 it decreased to 1.4 in the

strain carrying Erb1[R470E] (35% less), once again confirming that the mutation interfered with LSU assembly. We performed a western blot analysis of the resulting fractions in order to determine the localization of Nop7, Erb1 and Ytm1 in the context of wild-type Erb1 or Erb1[R470E] (Fig. 4.46).

Using anti-Erb1, anti-Nop7 and anti-Ytm1 antibodies we were able to detect the three proteins in the extracts. In the strain expressing Erb1, the proteins were mainly found in those fractions that sediment lower than 60S subunit and most likely correspond to pre-ribosomal particles with MW slightly higher than LSU. When compared to Erb1, in the context of Erb1[R470E], Nop7 did not show changes in its localization. Erb1[R470E] variant clearly tended to accumulate as free protein or was forming low MW complexes and sedimented with the pre-60S particles only to

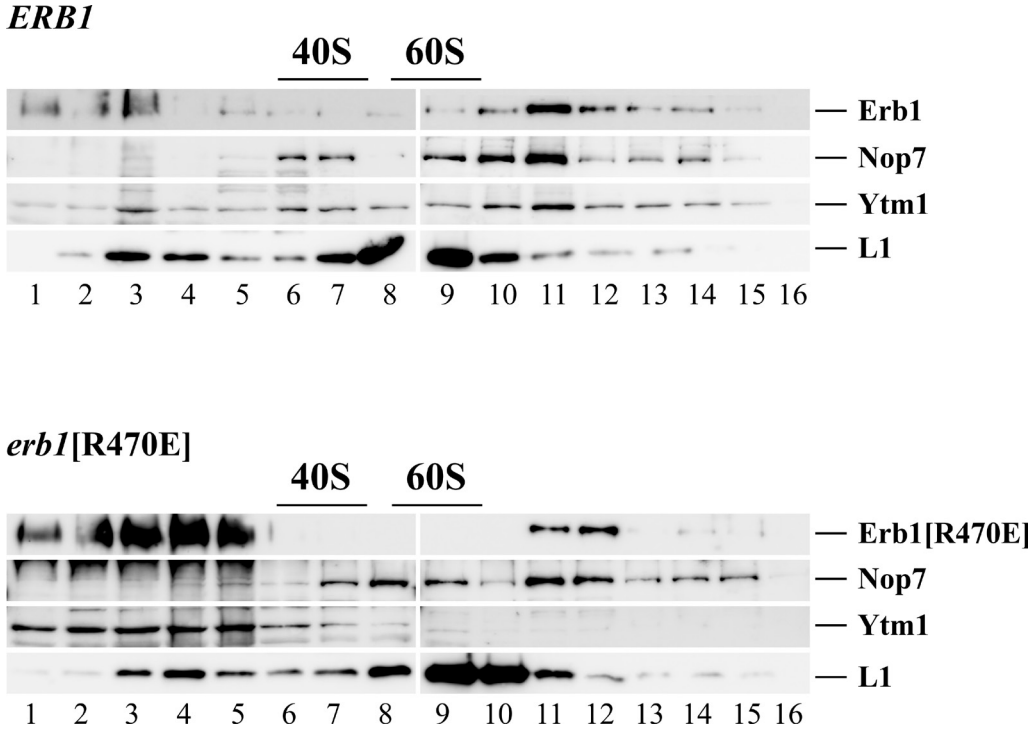
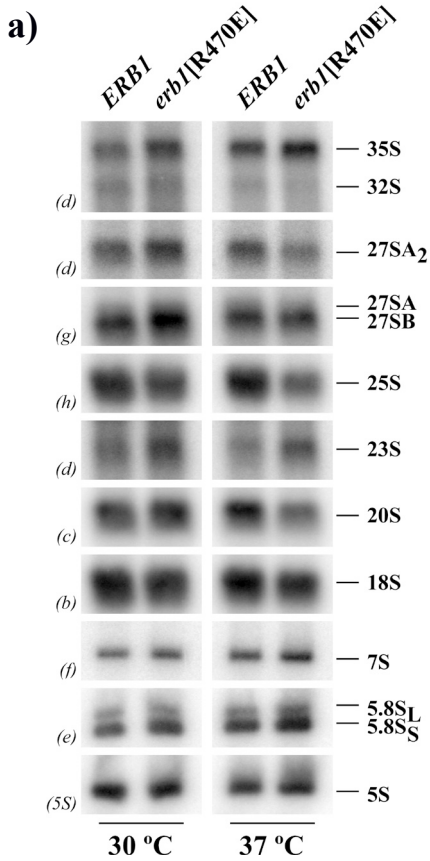


Figure 4.46. Erb1[R470E] disrupts the association of Ytm1 with pre-ribosomal particles. Ribosomal subunits were separated by ultra-centrifugation on sucrose gradients and collected fractions were subjected to Western Blot. In the mutant, higher amounts of Erb1[R470E] appeared as free protein not associated with pre-ribosomal fractions and almost all Ytm1 was detected at the top of the gradients dissociated from pre-60S. Ribosomal L1 protein was included as a control.

a lesser extent suggesting that the association between pre-ribosomes and mutant protein is weakened. Strikingly, Ytm1 protein was found only at the top of the gradient indicating that Erb1[R470E] drastically disrupted the recruitment of Ytm1 to the pre-60S.

Finally, we performed a northern blot analysis of the rRNA species detectable in wild-type or Erb1[R470] expressing strains. As in previous experiments, the effect of the mutation was more prominent at higher temperature. Erb1[R470E] led to a weak accumulation of 35S rRNA with concomitant reduction of $27SA_2$, 25S and 20S species (Fig. 4.47). This pattern of altered rRNA processing resembles the one previously described for other proteins that act at the same stage of ribosome biogenesis as Erb1 (A_3 factors).



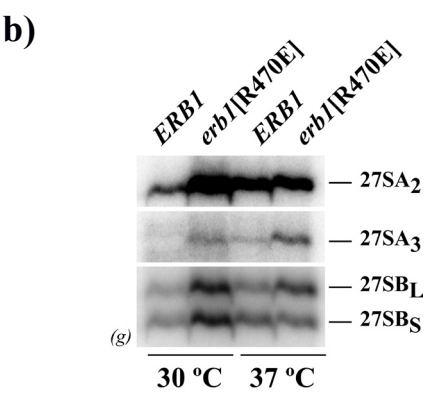


Figure 4.47. Mild effect of Erb1[R470E] in rRNA processing can be observed at 37ºC.
a) Northern blot showed that at 30ºC no significant difference between Erb1 and Erb1[R470E] could be observed. At 37ºC slight accumulation of 35S and decreased levels of 25S and 20S were detectable.
b) Primer extension showed that $27SA_3$ is accumulated and $27SB_S$ is reduced in Erb1[R470E] mutant.

5. DISCUSSION

Nop7, Erb1 and Ytm1 proteins bind to each other in order to assure the correct 60S assembly. The role of the trimer in the process is still poorly understood but it has been previously demonstrated that it is essential for cell viability in yeast and mammals. We have focused our efforts on the structural and biophysical characterization of the complex and then confirmed our findings using *Saccharomyces cerevisiae* as a model organism. Because many results in this work were obtained using proteins from *S. cerevisiae* or *Chaetomium thermophilum,* in this section, I will refer to the proteins as Erb1, Nop7 and Ytm1 indistinctly, unless otherwise specified.

5.1. Biophysical and biochemical characterization of the proteins

5.1.1. Production of Ytm1 requires eukaryotic expression system.

After our laborious attempts to over-express soluble Ytm1 in *E. coli* resulted unsuccessful, we managed to produce a highly homogeneous protein using insect cells. Since we could obtain UBL domain alone using bacteria, it became clear that the proper folding of the β-propeller of Ytm1 was the limiting step. Interestingly, it has been known that the seven-bladed WD40 domain of the β subunit of the trimeric G-protein relies on a eukaryotic chaperone CCT to acquire its final conformation (Kubota et al., 2006). It is tempting to think that a similar scenario takes place in the case of Ytm1. In addition, Rsa4 protein, which also contains an UBL and a C-terminal β-propeller domain and is structurally and functionally related to Ytm1, is also insoluble when produced in prokaryotic expression system (Bassler et al., 2014).

A high-throughput study searched for proteins that were binding co-translationally to SSB chaperone in yeast and both, Ytm1 and Rsa4, were identified (Willmund et al., 2013). SSB is known to facilitate correct folding of proteins that present slower translation rates. Moreover, co-chaperones of SSB and SSA called Zuo1 and Jjj1, respectively, that are known to stimulate its ATPase activity, have been shown to act during ribosome biogenesis (Albanèse et al., 2010; Meyer et al., 2007). Actually, yeast strains lacking Zuo1 or Jjj1 exhibit defects in 27S RNA processing, as when Ytm1 or its binding partners are depleted (Kaschner et al., 2015). Furthermore, Jjj1 is involved in the recycling of a ribosome assembly factor Arx1 through its interaction with Rei1 (Kressler et al., 2010; Ohmayer et al., 2013).

Zuo1 or Jjj1 might be able to recognize the nascent Ytm1 and/or Rsa4 polypeptide and bind to it in order to promote its folding. The UBL domain of both proteins might be involved in this kind of recognition because there are examples of chaperones that can selectively bind ubiquitin (Parcellier et al., 2003; Xu et al., 2013). As a matter of fact, Zuo1 can interact with ubiquitin as demonstrated in a high-throughput analysis (Phillips et al., 2013). The ability of Zuo1 or Jjj1 to assist folding of pre-ribosomal factors and maybe their association to pre60S still needs to be investigated.

A different possibility regarding the folding of Ytm1 and Rsa4 should also be taken into consideration. It has been proposed that the amino-terminal UBL domain could act as a chaperone that enhances the expression and proper spatial organization of the fused polypeptide (Karbstein, 2010). It has been shown that in several cases fusion of ubiquitin to certain proteins improved their heterologous expression and solubility in *E. coli* (Catanzariti et al., 2004). Therefore, it is possible that the N-terminal domain of Ytm1 assists the folding of the β-propeller but considering that ubiquitin and UBL alone can be expressed correctly in bacteria, this chaperoning activity would be expected to promote soluble expression of the full-length Ytm1 and, as shown by our expression assays, it does not. Thus it is highly probable that even if UBL makes a contribution to Ytm1 folding there should be additional mechanism, found exclusively in eukaryotes, that guarantees the expression of the functional protein.

5.1.2. Proteins from *Chaetomium thermophilum* are more stable *in vitro*

Upon expression of the proteins of interest we were able to compare the stability of Ytm1 and Nop7 from yeast with their counterparts from the thermophilic fungus. As expected, in both cases, proteins from *C. thermophilum* resulted to be more thermostable. Whereas, the mean Tm for Ytm1/Nop7 from yeast was 45ºC approximately, the thermophilic macromolecules presented Tm values close to 58-60ºC. This important shift in melting temperature had a beneficial effect on the expression levels and could positively affect crystallization of the macromolecules (Canaves et al., 2004). Indeed, we were able to obtain crystals of ChErb1-ChYtm1 and not of their counterparts from yeast. It is not easy to clearly explain the reasons behind the great thermal stability improvement in case of Ytm1 and Nop7. One

obvious difference between the proteins is their length; Ytm1 from yeast (51.4 kDa) has 35 residues less than ChYtm1 (52.9 kDa) and Nop7 (69.9 kDa) is 74 residues shorter than ChNop7 (77.1 kDa). In both cases the additional residues in the thermophilic proteins tend to accumulate in certain segments of the polypetide and, as seen in ChYtm1 structure, form extra motifs on the surface that could somehow stabilize the fold of adjacent regions or occlude flexible, less stable parts of the protein (Fig. 5.1). However some of those insertions are also present in other groups of Fungi that are not necessarily thermophilic, thus it is difficult to unambiguously explain the contribution of those sequences to the thermal stability of the protein.

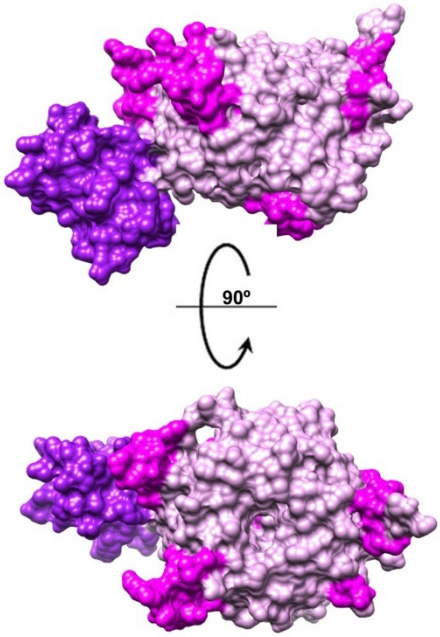


Figure 5.1 Additional elements within β-propeller of ChYtm1 appear on the surface of the domain.
All inserted loops (magenta) form protrusions on the surface of the β-propeller (pink) and do not distort the overall fold of the domain. UBL is shown in purple

Curiously, Erb1 from yeast (91.7 kDa) does not follow the rule established for Nop7/Ytm1 and is 8 residues longer than ChErb1 (90.7 kDa). The thermophilic protein does contain two extra segments within its N-terminal domain but in the C-terminal β-propeller it loses an important fragment of the insertion that occurs within WD2 as well as a part of the loop that connects blades 1 and 2. In this case, this loss of residues might not be related to the enhanced thermal stability but just to a lack of function and removal of dispensable segments during evolution, as discussed later.

5.1.3. Intrinsic lack of stability of Erb1 promotes crystallization of the β-propeller domain.

Previous studies on Erb1 in vivo have proposed that the levels of the protein could be an important point of control during ribosome biogenesis (Rohrmoser et al., 2007). In this work we show that Erb1 produced in vitro lacks stability and suffers sever proteolytic degradation in the crystallization drop, even in presence of its binding partners, Nop7 and Ytm1, suggesting that none of them bind to the cleaved area. There is a predicted PEST sequence in the central region of Erb1, between the conserved N-terminus and the β-propeller domain. PEST motifs are known to appear in proteins that undergo fast turn-over thus allowing a tight control of their levels in the cell. The PEST sequence and its effect on protein stability have been extensively studied in Eukaryotes. Although it has not been found in Bacteria, high content of proline - glutamic acid – serine - threonine can significantly increase the flexibility and disorder of the region which may make it easily accessible for proteases from *E. coli*. Random proteolysis events occurring during protein purification and crystallization yielded crystals of the β-propeller domain alone and in complex with Ytm1. High protein concentration used in crystallization trials probably accelerated proteolytic cleavage because during protein over-expression and purification Erb1 degraded only to a lesser extent. We could observe that, once concentrated up to 40 mg/ml and incubated for 24h at 4ºC, the protein degradation was much more prominent. Moreover, when PMSF 0.5mM was added to the sample used in the crystallization screens, we could not detect any crystal growth in the conditions that originally yielded crystals of truncated Erb1.

Another interesting issue regarding the proteolysis is the fact that upon addition of ChNop7, the dimer containing ChYtm1 and $ChErb1_{432-801}$ formed crystals with different morphology, packing and diffracting to a higher resolution. In yeast, Nop7 has been proposed to bind Erb1 through a central region of the protein (265-383). The same region in ChErb1 corresponds to residues 280-396. It is plausible that in presence of ChNop7 the proteolytic cleavage occurred in a different point, closer to the β-propeller domain, than when ChNop7 was absent. As a result, the carboxy-terminal fragment of ChErb1 remained flanked by a shorter peptidic fragment on its N-terminus that could pack in a different manner than in the

hexagonal crystals, where the N-terminus of the propeller points toward a gap between the symmetry related chains and is not traceable.

5.1.4. The β-propeller domain of Erb1 binds nucleic acids.

Previous studies have demonstrated that Erb1 could be UV cross-linked to rRNA within pre-ribosomal particles. Once we solved the structure of $Erb1_{416-807}$ we realized that it contained a large positively charged patch on the circumference of the domain. The size of the area and its degree of conservation led us to investigate whether it could be important from the functional point of view. Manual inspection in Coot confirmed that it was large enough to bind single-stranded RNA or DNA. We have also used HADDOCK to computationally check if there was enough space on the surface of the propeller, in the basic region, to accommodate a nucleic acid chain. We then decided to check if the interaction could be detected in vitro. We used the β-propeller domain of ChErb1 because it resulted to be more stable and expressed in higher amounts than its counterpart from yeast. The binding to unspecific RNA was confirmed by poly(U)-agarose beads, fluorescence and biolayer interferometry. The latter allowed us to calculate the binding affinity that resulted to be in a high nanomolar range (~170nM). Although we produced a set of ChErb1 variants containing single mutations of the conserved basic residues to alanine, we were not able to disrupt its interaction with RNA, probably because a more prominent modification of the patch is required to destabilize the binding.

Moreover, electromobility shift assay showed that ChErb1Ct was also able to associate to a fragment of dsDNA causing its retardation in agarose gel. Nevertheless, when we measured the affinity of the interaction between $ChErb1_{432-801}$ and biotynilated-ssDNA, it resulted to be lower than in the case of ssRNA. In addition, Nop7, Erb1 and truncated fragments of both, tend to co-purify with large amounts of nucleic acids from *E. coli* giving further evidence that the binding takes place between the proteins and RNA and/or DNA. Given the presence of the complex within pre60S and its higher affinity for RNA, we consider that the binding to DNA might not be relevant from the functional point of view and could occur due to electrostatic attraction between negatively charged nucleic acid and the proteins.

We have not measured the interaction affinity between Erb1Ct and the specific rRNA sequence described by Grannenman et al but we prove that the β-

propeller can bind RNA regardless of its sequence. The specificity of the binding in the context of pre60S could be provided by Nop7 that is required for the recruitment of Erb1 to the pre-ribosomal complex. The ability of WD40 domain of Erb1 to bind RNA could play secondary role and stabilize the association.

5.1.5. Erb1 oligomerization in solution

When members of the PeBoW complex from *S. cerevisiae* and *C. thermophilum* were expressed and purified we realized that Erb1, and to a lesser extent Nop7-Erb1-Ytm1, formed higher MW oligomers in gel filtration.

In case of Erb1, it tends to form dimers and higher MW oligomers in a concentration-dependent manner. Gel filtration pattern clearly shows peaks that correspond to oligomers and we have confirmed this behavior by native gel electrophoresis. Moreover, when the monomer-containing sample is concentrated and re-injected into Superdex column, oligomeric species appear. The lack of homogeneity in the sample has a negative effect on the crystallization of Erb1 and might explain why the full-length protein did not crystallize when the random proteolysis was inhibited. Moreover, we have not been able to measure the binding affinity between Erb1 and Nop7 by ITC probably due to a highly dynamic equilibrium occurring between the different oligomers in solution. Since we have seen that when Erb1 is incubated with Nop7 and Ytm1 a heterotrimeric species is predominant, we believe that Erb1-Erb1 interaction is less favorable in the context of the complex. Still, as seen in gel filtration, Nop7-Erb1-Ytm1 particle also presents oligomerization states in which the 1:1:1 ratio is altered. Given the ability of Erb1 to dimerize it is logical to think that it could induce the formation of hexamers, or even larger complexes, composed of Nop7-Erb1-Ytm1 proteins in vitro. As seen in gel filtration and crystal structure of $Erb1_{416-807}$ the β-propeller domain of Erb1 does not form dimers thus the propensity to oligomerize could be attributed to the N-terminal domain of the protein. It is also possible that the dimers of Erb1 are formed due to the electrostatic interactions, in a head-to-tail fashion, between the positively charged C-terminus (theoretical pI = 8.7) and highly acidic N-terminal region (pI = 4.5).

5.2. Structural analysis of Erb1, Ytm1 and their interaction

5.2.1. The sequence and structure of the β-propeller of Erb1 are well conserved

From a structural point of view the β-propeller of Erb1 highly resembles other seven-bladed WD40 domains. Although its sequence is well conserved within Erb1/Bop1 family, it is not very similar to other WD repeats-containing proteins. Initially it was difficult to unambiguously recognize seven WD repeats mainly due to the long insertion that appears within WD2. Actually this fragment is the most distinguishable feature in the structure of the domain. Additional alpha-helices and a β-strand form an important protrusion that alters the globularity of the WD40 fold. We observe an interesting rearrangement of the interactions between the blades 1-2 and 2-3 in order to maintain the stability of the domain. As a result, the insertion increases the area of the bottom part of the propeller that, as discussed later, is important for the interaction with Ytm1. In the context of the conservation, the sequence of the extension is the most variable part of the domain although when we compare the structure of Erb1 from *S. cerevisiae* with the one of *C. thermophilum*, they are not significantly different. It is important to bear in mind that the crystal structure of ChErb1Ct has been solved in presence of ChYtm1 and the binding of both macromolecules might have an effect on the relative position of the loop between strand "2c"and the α-helix 2. In any case, the superimposition of both domains shows that the overall architecture is perfectly preserved between both organisms. When we compare the sequence of the insertion itself, it clearly tends to be shorter in higher Eukaryotes (Fig. 5.2). As already mentioned, the flexible part of the β-propeller that could not be traced in our crystal structures is much shorter in $ChErb1_{432\text{-}801}$ than in $Erb1_{416\text{-}807}$ and if we compare it with higher eukaryotes, it practically disappears in Bop1 from mammals. Therefore it is possible that initially a long insertion appeared within WD2 and it has been sequentially shortened during evolution suggesting that the part that cannot be seen in the structure lacks functional relevance.

Another well conserved feature seen in the structure of Erb1Ct is the electrostatic potential. In both, $Erb1_{416\text{-}807}$ and $ChErb1_{432\text{-}801}$, we could observe a long breach of positively charged residues that presumably bind nucleic acids. The size and the shape of the area are large enough to bind ssRNA as shown by docking.

The low variability in this region of the propeller may, once again, indicate that the RNA binding is not just an artifact and could have an important role in Erb1 association with pre-ribosomes

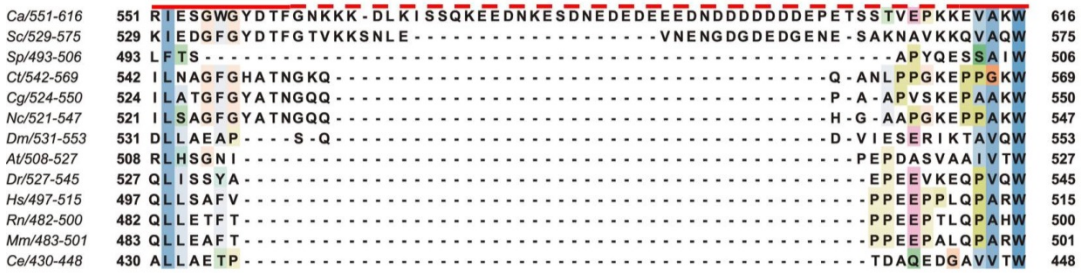


Figure 5.2 The insertion within WD2 of Erb1 disappears during evolution.
The part of the insertion that could nor be traced in our models (dashed red line) is much shorter in higher eukaryotes. Conserved residues are colored. Ca: *Candida albicans*, Sc: *Saccharomyces cerevisiae*, Sp: *Schizosaccharomyces pombe*, Ct: *Chaetomium thermophilum*, Cg: *Chaetomium globosum,* Nc: *Neurospora crassa*, Dm: *Drosophila melanogaster*, At: *Arabidopsis thaliana*, Dr: *Danio rerio;* Hs: *Homo sapiens*, Rn: *Rattus norvegicus,* Mm: *Mus musculus*, Ce: *Caenorhabditis elegans*

5.2.2. The structure of Ytm1 manifests that highly variable sequence can render perfectly conserved folds.

The structure of Ytm1 from *Chaetomium thermophilum* confirmed the predicted architecture of the protein. The N-terminal ubiquitin-like domain is followed by a seven-bladed β-propeller and both domains are essential for the function. The WD40 domain in case of Ytm1 is more unusual than the one of Erb1. It means that even though the core of the propeller superimposes well with other WD40 folds, it is decorated by many insertions and extended loops that adopt secondary structures on the surface of the domain. While some of these extra motifs can be related to certain groups of fungi, the rest seems to be evolutionary conserved and also appear in mammalian Wdr12 proteins. Unsurprisingly, all inserted segments appear on the surface of the protein, in different parts of the propeller (as already shown in Fig. 5.1). Interestingly, none of those extra residues are found on the top face of the β-propeller domain, which is the hot-spot area involved in the interaction with Erb1. One of the insertions forms a β-strand, one folds into an α-helix and the other three are parts of more less extended loops between the strands of the

propeller. The presence of additional elements in the β-propeller structure is not surprising taking into account the great number of eukaryotic proteins that carry similar domains. The most known function of WD40 repeats is to mediate protein-protein interactions and it is clear that there must exist a way to provide specificity to those large folds. Whereas, the canonical arrangement of the β-sheets gives rise to the typical propeller, the loops and extra helices can play an important role in the selective binding to other macromolecules.

Curiously, the insertion that appears between blades 6 and 7 resembles a similar segment in another β-propeller, the one of Asc1 (Rack1 in higher eukaryotes) (Coyle et al., 2009). It has been described to form a knob structure that appears in the interface that is involved in binding of Asc1 to SSU in order to control translation (Fig. 5.3a). For the crystal structure of Asc1 it has been shown that this fold is maintained due to a network of hydrophobic interactions. In case of ChYtm1 we can observe an important salt bridge between E442 and R454 (Fig. 5.3b).

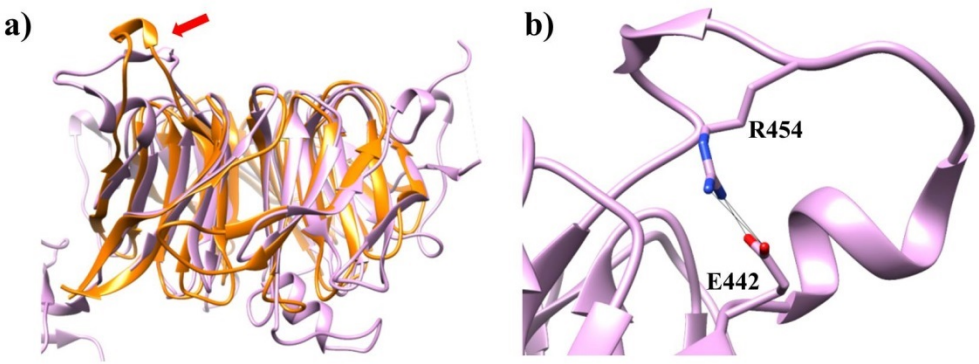


Figure 5.3 A knob-like element between blades 6 and 7 also appears in Rack1.
a) The insertion between blades 6 and 7 forms a knob-like structure (red arrow) on the top face of the β-propeller of ChYtm1 (pink) that has been described in Rack1 protein from yeast (orange, PDB:3frx). b) The knob of ChYtm1 is maintained mainly by a salt bridge between E442 and R454 although these residues are not conserved.

Since the deletion of the knob did not impair binding of Asc1 to the ribosome, it has been proposed that it could be important for a specific recognition of other ligands, especially because it lies close to the hot-spot top face of WD40 domain. This is the case of ChYtm1 because the knob does establish a series of

contacts with ChErb1Ct and contributes to the binding. It is noteworthy that two proteins that carry a seven-bladed β-propeller and bind to ribosome (Ytm1 to pre-60S and Rack1 to 40S) share a very characteristic element in the same position within their tertiary structure.

In the crystal structure of ChYtm1 some of the extended loops that appear in the β-propeller could not be modeled with high accuracy or present high B-factors due to their flexibility. Given the clear degradation pattern of Ytm1 upon expression carried out for more than 72h, it is also possible that some of those fragments suffer proteolytic cleavage that does not disrupt the propeller structure but explains the lack of electron density in several regions of the asymmetric unit.

One of the parts that could not be traced in our model is the segment between V272 and S283. It belongs to a long fragment that separates WD3 and WD4. Unlike the insertion from Erb1Ct, this sequence of ChYtm1 contains several fully conserved amino acids (Fig. 5.4a).

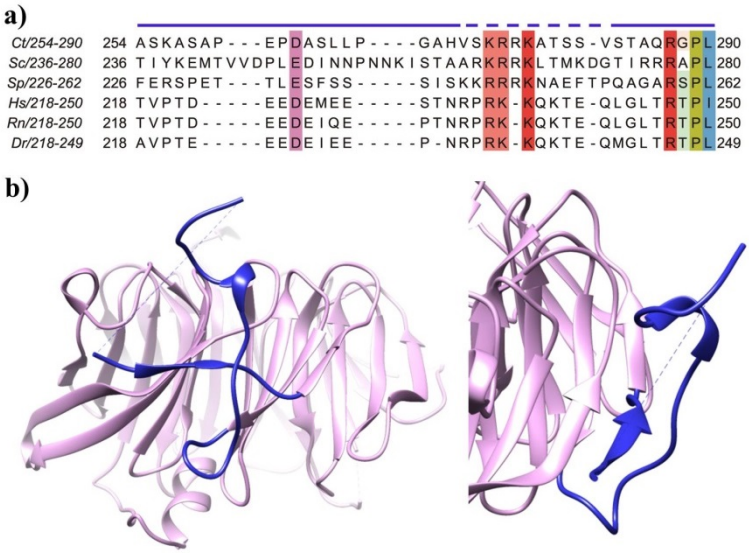


Figure 5.4 Partially conserved insertion between WD3 and WD4 of ChYtm1 forms a cross-like fold. a) Alignment of the sequences corresponding to the region that separates WD3 and WD4 shows that few residues are fully conserved (shown in color). It contains an NLS that overlaps with the region that is not visible in the crystal (dashed blue line). b) The inserted fragment (blue) forms a long loop parallel to the axis of the propeller and perpendicular to the 5th strand of blade 2.

The part of the insertion that we traced folds in a very peculiar way. It forms a cross-like arrangement because first it goes from the strand "3c" out of the core of the propeller forming a long loop and then it becomes the fifth β-strand of the blade

2. Finally it goes through the gap between the propeller and the loop until it reaches the strand "4d", this way tiding the knot (Fig. 5.4b).

On one hand, the uppermost part of the insertion is involved in binding to ChErb1 (box C in Fig. 4.38). On the other hand, the missing fragment of the structure contains a potential Nuclear Localization Sequence (NLS) which probably is important for the nuclear import of Ytm1. Given the localization of the insertion on the surface, its characteristic fold might be recognized by importins or other factors that could translocate Ytm1 to the nucleus. This is in line with the fact that shuttle of Ytm1 to the nucleus is independent of Nop7-Erb1.

Similarly to the β-propeller, ubiquitin-like domain of Ytm1 constitutes another good example of a domain whose structure is well preserved in spite of a poor sequence conservation. It is remarkable how the whole ubiquitin-like fold can be assembled even though the composition of the N-terminal part of Ytm1 has very little in common with ubiquitin. Even when we compare UBL from Ytm1 with the same domain of Rsa4, we observe only a small number of residues that match. From the functional point of view, the most documented role of UBL from Ytm1 is its binding to the MIDAS domain of Rea1. The crystal structure shows that E88, the residue described to be fundamental for this interaction to occur, appears in a loop that acquires similar orientation as the one from Rsa4 that also is known to bind Rea1. The interaction of Ytm1 with AAA-ATPase has been described as necessary to remove Nop7-Erb1-Ytm1 from pre60S (at least a partial release of Nop7 has been observed). It has been proposed that the energy obtained from the hydrolysis of ATP is then transformed into a mechanochemical force that, through UBL domain of Ytm1, induces conformational changes required for the dissociation of the factors. The fact that we found two different orientations of UBL in two different crystals of ChYtm1-ChErb1 complex is line with this hypothesis and confirms that a certain degree of flexibility is allowed between the N-terminal and the C-terminal domains of Ytm1. Since similar observations had been made previously in the case of Rsa4, our findings once again prove functional parallelism regarding both proteins.

5.2.3. Erb1-Ytm1 dimer is maintained through a large interface between their C-terminal domains.

Undoubtedly, the most remarkable finding reported in this work is the fact that the β-propeller domain of Erb1 binds to Ytm1 protein. In recent years it has been postulated that the carboxy-terminal region of Erb1 was not involved in ribosome biogenesis and no other evident function has been proposed. When we first solved the structure of the β-propeller alone, we showed that it could bind RNA and that it had a very conserved area on the surface that could be involved in association to other molecules. In order to prevent Erb1 degradation through its central fragment, which was believed to bind Ytm1, we re-constituted the trimer and surprisingly still observed proteolysis. When crystallization trials of ChNop7-ChErb1-ChYtm1 yielded good quality crystals, we surprisingly discovered that ChYtm1 appeared bound to the β-propeller domain of ChErb1. Taking into account previous reports regarding the interactions required for Nop7-Erb1-Ytm1 complex formation, we carefully investigated if the content of the asymmetric unit was not an artifact induced by the crystal. We were able to confirm the interaction *in vitro* and we proved that that both macromolecules associated with a good affinity (K_D ~ 5nM). We further characterized the interaction *in vitro* and showed that there was no need for the N-terminal domain of Erb1 to facilitate or enhance the assembly of ChErb1-ChYtm1 dimer.

Both β-propeller domains bind in a unique manner that has not been described previously for WD40-WD40 interaction. Even though the β-propellers are large platforms that mediate protein-protein contacts, there are not many examples of structures where two of them directly interact (Mylona et al., 2006). In the case of ChYtm1-ChErb1$_{432-801}$ dimer we observe that the top face of ChYtm1 is attached to the lower part of the β-propeller of ChErb1 which result in approximately 55º angle between the central axes of both domains.

We have learned from the structure that the binding takes place through a large surface comprising the top face of the β-propeller of ChYtm1 and the bottom part of ChErb1. The hot-spot area surrounding the central channel of ChYtm1 recognizes a platform created mostly by blades 1, 2 and 7 as well as by the insertion attached to blade 2 from ChErb1. From a functional point of view we should

remember that it has been proposed that Erb1 is recruited to the pre-ribosome before it interacts with Ytm1, thus the area of contact between both β-propellers should be exposed and ready for the docking of Ytm1. There may exist a mechanisms of regulation of such an interaction in the sense that when Nop7-Erb1 resides in a pre60S particle that has not yet reached maturation level at which the complex should be released, the area of Erb1 that is recognized by Ytm1 might be occluded or present modifications that would impede strong interaction with Ytm1. Alternatively, Ytm1 could bind shortly after loading of Erb1 on the pre-ribosome and the liberation of Nop7-Erb1-Ytm1 could depend on the interplay between Ytm1 and Rea1. Both mechanisms are not mutually exclusive and may occur at two sequential stages of ribosome maturation.

5.2.4. Mutational analysis of ChErb1-ChYtm1 binding

After a detailed manual and computational analysis of the binding interface we have selected those residues that seemed to have major impact on the interaction. We prepared a set of ChErb1 variants that contained single point mutations in the most conserved region involved in the association. We investigated the impact of those mutations on the binding and we identified that a single substitution R486E was able to decrease the affinity of the interaction by two orders of magnitude (K_D ~ 249nM). It was an important finding because we proved that altering only one salt-bridge we could strongly affect the interaction between two large protein surfaces. Moreover, R486E mutant manifested the relevance of the high degree of conservation that we had observed within ChErb1 area (residues 481-487) that contacts ChYtm1 through its central channel.

We crystallized the dimer of ChYtm1 and $ChErb1_{432-801}$[R486E] in order to prove that the mutation did not destabilize the domain and, at the same time, we confirmed that thermal stability of ChErb1R486E β-propeller was not reduced. This observation is relevant because any mutation that alters the architecture of WD40 domain could exert an indirect negative effect on the association to Ytm1. This situation has been described by Miles et al. (Miles et al., 2005) who identified a temperature-sensitive mutant of Ytm1 in yeast, *ytm1-1,* that contained G398D and S442N mutations. Those residues correspond to G411 and S475 in *C. thermophilum*, respectively. Ytm1-1 could not stably associate with Erb1 in vitro and it affected the

stability of Nop7-Erb1 dimer in yeast. Although, as seen in ChYtm1-ChErb1Ct structure, G411 and S475 residues are not directly involved in binding to ChErb1, they are located close to the interface between blades 6 and 7 of ChYtm1 that interacts with blade 7 of ChErb1. Moreover, S475 contributes to the blade stability through a network of hydrogen bonds with neighboring residues. Thus, it is possible that the mutations in this region affect proper folding of blades 6 and 7 and perturb the area of interaction with ChErb1.

5.3. Re-defining the role of the β-propeller domain of Erb1 in *Saccharomyces cerevisiae*

The structural information gathered from the analysis of Erb1-Ytm1 dimer from *C. thermophilum* was then translated into functional studies *in vivo* using Baker's yeast. We propose that if the β-propeller domain of Erb1 does bind Ytm1 in the cell, a mutation within the domain that decreases the biding affinity should affect negatively the association between both proteins. Secondly, if the interaction between Erb1 and Ytm1 is relevant during LSU assembly, we should observe an effect if that association is altered *in vivo*. Since ChErb1[R486E] resulted to have lower affinity for ChYtm1 than the wild-type protein we decided to study if the corresponding mutation in Erb1 from yeast (R470E) would have any effect *in vivo*.

To confirm our hypothesis we used a diploid strain erb1/Δerb1 that allowed us to generate a Δerb1 strain complemented with YCplac33-Erb1. We then checked if YCplac111-Erb1 or YCplac111-Erb1[R470E] were able to complement the depletion of Erb1 in Δerb1 strain upon loss of YCplac33-Erb1 in a medium containing 5-FOA. As expected, the wild type copy of Erb1 from Ycplac111 could rescue the lethal phenotype of Erb1 depletion at 30ºC and 37ºC. Whereas Erb1[R470E] - complemented strain grew nearly as well as the wild-type at 30ºC, at 37ºC it clearly showed slower growth in YPD. These findings led us to further investigate the effect of the mutation in the cell. The most obvious reason for slower growth in the context of Erb1[R470E] could be altered ribosome biogenesis due to defective binding of Ytm1 and Erb1[R470E]. We performed polysome profiling and we could observe an effect of the mutations that manifested in reduced levels of 60S subunit and aberrant translation. Once again, the phenotype was stronger at 37ºC and confirmed that LSU assembly was not fully functional. Actually, the polysome

profiles were strikingly similar to the one seen previously for *ytm1-1* mutant that also affects the interaction between Ytm1 and Erb1. Upon separation of ribosomal subunits on sucrose gradients we performed a western blot of the collected fractions and we observed that Erb1 wild-type, Nop7 and Ytm1 mostly appeared associated to the pre-60S. When we analyzed the strain expressing Erb1[R470E] Nop7 presented the same pattern as in the case of the wild-type starin. Erb1[R470E], however, was predominantly found as free protein because its stable association with pre60S was less efficient. More interestingly, the binding of Ytm1 to the pre-ribosomal particles was almost completely disrupted. This observation is in line with previous reports regarding the stability of Nop7-Erb1-Ytm1 trimer. It has been shown that, in the context of ytm1-1 mutation the binding of Erb1 to the pre-60S is less efficient. Similar situation may take place when the binding alteration is induced by a mutation in Erb1 (R470E). Even though Erb1[R470E] can still bind Nop7 as well as the wild-type protein and is efficiently recruited to the pre-ribosome, this association is more transient because the binding of Ytm1 is not effective. In this scenario the stability of the whole trimer is diminished. Moreover, for obvious reasons, if Ytm1 and Erb1 do not interact properly, Rea1-dependent removal of Nop7-Erb1-Ytm1 from pre60S is defective. Altogether, these phenomena result in abnormal, less efficient ribosome production that reduces translation rates which consequently delays cell growth. Because we observed that the β-propeller domain that carries R486E mutation was properly folded and as stable as the wild-type domain, it was highly unlikely that the effect of Erb1[R470E] in yeast was caused by misfolded protein.

This was an important proof of concept because previous functional studies in yeast seemed to indicate that the total lack of the C-terminal domain of Erb1 would not exhibit any effect on ribosome biogenesis or cell growth. Given our structural findings which suggested that the β-propeller of Erb1 was relevant for Nop7-Erb1-Ytm1 complex formation, we hypothesized that if a mutation within the domain interfered with the assembly of the trimer, we should be able to observe its consequences. On the other hand, if the β-propeller had no role in binding to Ytm1 or in ribosome biogenesis in general, mutations of its amino acids should not cause any irregularities. Another possibility to consider was the scenario in which R470E mutation would not be sufficient to alter Erb1-Ytm1 binding in vivo.

5.4. Summary of the interactions that hold Nop7-Erb1-Ytm1 complex together

Our biophysical, structural and in vivo analysis has allowed us to describe the way Nop7, Erb1 and Ytm1 interact. Nop7 uses its N-terminal domain to interact with the central region of Erb1. Nor the BRCT domain of Nop7, nor its two putative coiled-coils are required for this binding. From previous studies and from the results of this work, it seems obvious that the amino-terminus that harbors well conserved PES1-Nt domain of unknown architecture is responsible for the association between both factors to occur (Fig. 5.5).

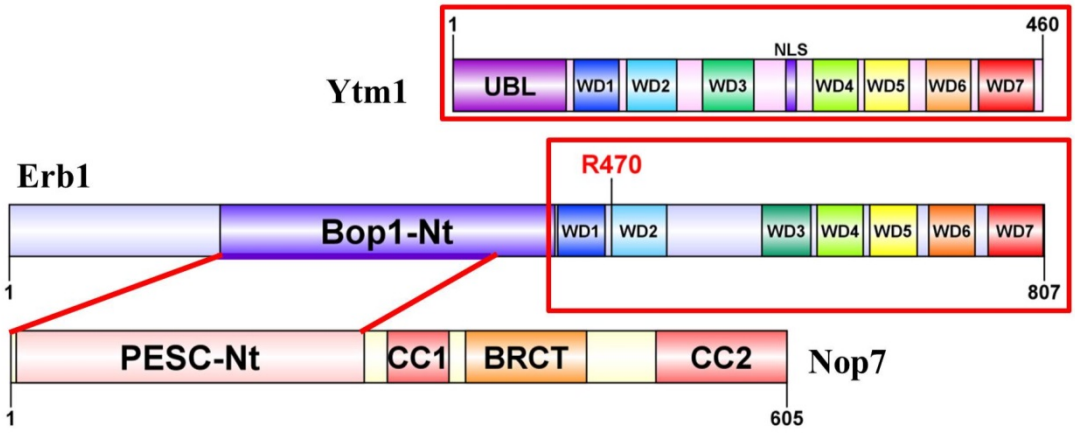


Figure 5.5 Summary of the interactions between Nop7, Erb1 and Ytm1.
Diagram representing Nop1, Erb1 and Ytm1. Domain organization of each protein is shown. The area of interaction between Nop7 and Erb1 is indicated with red lines. Segments of Ytm1 and Erb1 that interact and have been crystallized and described in this work are marked with red rectangles. The position of the arginine that has been modified in Erb1[R470E] mutant is shown and labeled.

Regarding the C-terminal domain of Erb1, altogether, our results clearly indicate that this large β-propeller may act as an important platform that binds several macromolecules within pre60S particle. When we represent at the same time the Ytm1 binding area and the positive patch on the surface, they localize on the opposite sides of the domain (Fig. 5.6). Therefore, simultaneous interaction with both, Ytm1 and rRNA is perfectly compatible. Moreover, in several occasions we have seen that upon binding to Ytm1, Erb1Ct can still bind nucleic acids. Interestingly, recent studies regarding Drs1 helicase and its mammalian counterpart Ddx27 have shown that it interacts with Erb1. There is not enough evidence to

confirm that both proteins associate physically because they bind RNA, which could mediate some sort of indirect interaction. Nevertheless, it has been shown that the N-terminally truncated Bop1 (mammalian orthologue of Erb1) still interacts with Drs1 in total cell extracts, whereas the same protein lacking the β-propeller domain does not. These findings seem to suggest that the C-terminal half of Erb1, once again, might be responsible for a specific macromolecular interaction. Further investigation should be carried out to address this interesting hypothesis.

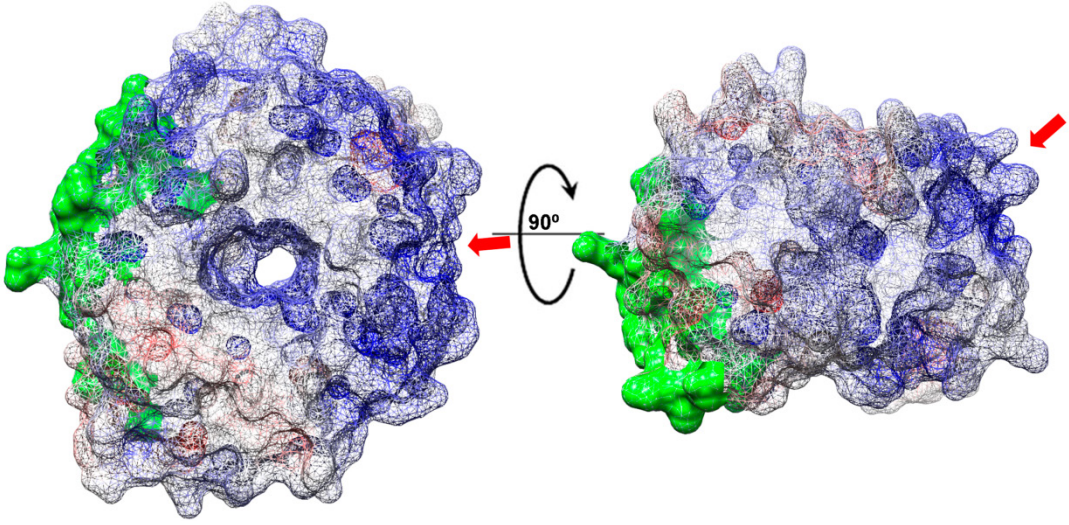


Figure 5.6 Electrostatic surafce view of the β-propeller of Erb1.
Electrostatic surface potential of ChErb1Ct represented as mesh shows that highly electropositive area possibly involved in binding to RNA (red arrow) is located on the opposite side to ChYtm1 binding site (in green)

Although the exact role of Nop7-Erb1-Ytm1 trimer in ribosome assembly is not well understood, it has been proposed to act as a scaffold necessary for the structural reorganization of the nascent ribonucleoprotein. If we consider that Erb1 and Nop7 might interact with Drs1 modulating its function on the pre-ribosome, it is plausible that upon initial binding of Nop7-Erb1 to pre60S, the dimer performs a Drs1-related task, which once accomplished allows the recognition of Erb1 by Ytm1 and the liberation of the trimer from the machinery. The removal of Nop7-Erb1-Ytm1 seems to be important for the assembly to continue; therefore it is likely to induce important conformational changes within emergent LSU. Further

investigations need to be carried out in order to shed light on the exact role of Nop7 sub-complex, its architecture and the regulation of its assembly.

At last, we have proven that by affecting highly conserved interface between two proteins that are essential in ribosome biogenesis it is possible to alter cell growth. It is an interesting observation if we imagine how deleterious effects could have the disruption of ribosome synthesis in quickly proliferating cells. The area of the interaction that we targeted in our mutational analysis is invariable. However, as seen in the structure, there are some points of contact that are not preserved in higher eukaryotes. Thus, it would be interesting to assess the effect of a mutation analogous to R470E from yeast, in mammalian cells. If the disruption of the salt bridge between Bop1 and Wdr12 in mammals also diminishes their association rate and produces negative effect on the cell growth, it will open a whole new range of possibilities to target cell proliferation at ribosome biogenesis level.

The results of this work clearly demonstrate how the structural information provided by protein crystallography can lead to the functional re-evaluation of the protein complexes. At the same time it becomes obvious that biophysical and structural methods are a powerful tool to verify results obtained by genetic and functional studies.

6. CONCLUSIONS

1. The carboxy-terminal domain of Erb1 is a seven-bladed β-propeller that has the ability to bind nucleic acids *in vitro*. It can bind both, RNA and DNA, but manifests higher affinity for RNA.
2. Ytm1 protein requires a eukaryotic expression system for its proper folding and this requirement is caused by the C-terminal domain of the protein.
3. Ytm1 carries an ubiquitin-like domain (UBL) on its amino-terminus that possesses a well conserved glutamic acid residue (E88 in *C. thermophilum*) on an exposed loop that mediates its binding to the MIDAS domain of Rea1.
4. UBL domain is attached by a flexible linker to the carboxy-terminal segment of the protein, which folds into a seven-bladed β-propeller structure and contains a large number of additional motifs that significantly modify the surface of the domain.
5. Crystal structure of Ytm1-Erb1 dimer from *C. termophilum* shows that the top face of the β-propeller of Ytm1 binds with high affinity to the bottom part of the β-propeller of Erb1, through a large interface. The way both domains interact has not been previously seen for associating β-propellers.
6. The interaction between Ytm1 and Erb1 does not depend on the amino-terminal segment that precedes the β-propeller of Erb1, as suggested before.
7. A point mutation on the surface of the β-propeller of Erb1 (R486E) that disrupts the salt bridge between R486 and D112 (in *C. thermophilum*) decreases its affinity for Ytm1 by two orders of magnitude in vitro but does not affect structural or biophysical properties of the domain.
8. Erb1[R470E], a mutation equivalent to R486E from *C. thermophilum,* in *S. cerevisiae* does not rescue the lethal phenotype of Δerb1 to the wild-type extent and induces slow growth.
9. 60S subunit biogenesis and rRNA processing are affected in the context of Erb1[R470E] mutation in *S. cerevisiae*, as shown by polysome profiling and northern blot analysis.
10. Erb1[R470E] associates with pre-ribosomal particles less efficiently than Erb1 wild-type and in consequence higher levels of free Erb1[R470E] can be detected in yeast.
11. The β-propeller domain of Erb1 is involved in the stable recruitment of Ytm1 to the nascent pre-60S particle. Thus, Erb1[R470E] drastically reduces pre-ribosomal incorporation of Ytm1.

7. REFERENCES

Adams, C.C., Jakovljevic, J., Roman, J., Harnpicharnchai, P., Woolford, J.L., 2002. Saccharomyces cerevisiae nucleolar protein Nop7p is necessary for biogenesis of 60S ribosomal subunits. Rna 8, 150–165. doi:10.1017/S1355838202010026

Adams, P.D., Afonine, P. V, Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66, 213–21. doi:10.1107/S0907444909052925

Albanèse, V., Reissmann, S., Frydman, J., 2010. A ribosome-anchored chaperone network that facilitates eukaryotic ribosome biogenesis. J. Cell Biol. 189, 69–81. doi:10.1083/jcb.201001054

Albert, B., Perez-Fernandez, J., Léger-Silvestre, I., Gadal, O., 2012. Regulation of ribosomal RNA production by RNA polymerase I: does elongation come first? Genet. Res. Int. 2012, 276948. doi:10.1155/2012/276948

Amlacher, S., Sarges, P., Flemming, D., van Noort, V., Kunze, R., Devos, D.P., Arumugam, M., Bork, P., Hurt, E., 2011. Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. Cell 146, 277–89. doi:10.1016/j.cell.2011.06.039

Auende, M.L., Amsterdam, A., Becker, T., Kawakami, K., Gaiano, N., Hopkins, N., 1996. Insertional mutagenesis in zebrafish identifies two novel genes , pescadillo and dead eye , essential for embryonic development 3141–3155.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. 98, 10037–41. doi:10.1073/pnas.181342398

Barna, M., 2013. Ribosomes take control. Proc. Natl. Acad. Sci. U. S. A. 110, 9–10. doi:10.1073/pnas.1218764110

Bassler, J., Kallas, M., Pertschy, B., Ulbrich, C., Thoms, M., Hurt, E., Baßler, J., 2010. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. Mol. Cell 38, 712–21. doi:10.1016/j.molcel.2010.05.024

Bassler, J., Paternoga, H., Holdermann, I., Thoms, M., Granneman, S., Barrio-Garcia, C., Nyarko, A., Stier, G., Clark, S. a, Schraivogel, D., Kallas, M., Beckmann, R., Tollervey, D., Barbar, E., Sinning, I., Hurt, E., Baßler, J., 2014. A network of assembly factors is involved in remodeling rRNA elements during preribosome maturation. J. Cell Biol. 207, 481–98. doi:10.1083/jcb.201408111

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., Yusupov, M., 2011. The structure of the eukaryotic ribosome at 3.0 Å resolution. Science 334, 1524–9. doi:10.1126/science.1212642

Berg, J.M., Tymoczko, J.L., Stryer, L., 2002. Biochemistry, Fifth Edition. W.H. Freeman.

Bernstein, K.A., Granneman, S., Lee, A. V, Manickam, S., Baserga, S.J., 2006. Comprehensive Mutational Analysis of Yeast DEXD / H Box RNA Helicases Involved in Large Ribosomal Subunit Biogenesis 26, 1195–1208. doi:10.1128/MCB.26.4.1195

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., Boeke, J.D., 1998. Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14, 115–32. doi:10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2

Canaves, J.M., Page, R., Wilson, I.A., Stevens, R.C., 2004. Protein biophysical properties that correlate with crystallization success in Thermotoga maritima: maximum clustering strategy for structural genomics. J. Mol. Biol. 344, 977–91. doi:10.1016/j.jmb.2004.09.076

Catanzariti, A.-M., Soboleva, T.A., Jans, D.A., Board, P.G., Baker, R.T., 2004. An efficient system for high-level expression and easy purification of authentic recombinant proteins. Protein Sci. 13, 1331–9. doi:10.1110/ps.04618904

Chaudhuri, I., Söding, J., Lupas, A.N., 2008. Evolution of the beta-propeller fold. Proteins 71, 795–803. doi:10.1002/prot.21764

Chen, C.K.-M., Chan, N.-L., Wang, A.H.-J., 2011. The many blades of the β-propeller proteins: conserved but versatile. Trends Biochem. Sci. 36, 553–61. doi:10.1016/j.tibs.2011.07.004

Cheng, L., Li, J., Han, Y., Lin, J., Niu, C., Zhou, Z., Yuan, B., Huang, K., Li, J., Jiang, K., Zhang, H., Ding, L., Xu, X., Ye, Q., 2012. PES1 promotes breast cancer by differentially regulating ER α and ER β 122. doi:10.1172/JCI62676.In

Chevaillier, P., 1993. Pest sequences in nuclear proteins. Int. J. Biochem. 25, 479–482. doi:10.1016/0020-711X(93)90653-V

Ciganda, M., Williams, N., 2011. Eukaryotic 5S rRNA biogenesis. Wiley Interdiscip. Rev. RNA 2, 523–33. doi:10.1002/wrna.74

Clapperton, J.A., Manke, I.A., Lowery, D.M., Ho, T., Haire, L.F., Yaffe, M.B., Smerdon, S.J., 2004. Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. Nat. Struct. Mol. Biol. 11, 512–8. doi:10.1038/nsmb775

Coyle, S.M., Gilbert, W. V, Doudna, J. a, 2009. Direct link between RACK1 function and localization at the ribosome in vivo. Mol. Cell. Biol. 29, 1626–34. doi:10.1128/MCB.01718-08

De la Cruz, J., Lacombe, T., Deloche, O., Linder, P., Kressler, D., 2004. The putative RNA helicase Dbp6p functionally interacts with Rpl3p, Nop8p and the novel trans-acting Factor Rsa3p during biogenesis of 60S ribosomal subunits in Saccharomyces cerevisiae. Genetics 166, 1687–99.

De Vries, S.J., van Dijk, M., Bonvin, A.M.J.J., 2010. The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc. 5, 883–897. doi:10.1038/nprot.2010.32

Deana, A., 2005. Lost in translation: the influence of ribosomes on bacterial mRNA decay. Genes Dev. 19, 2526–2533. doi:10.1101/gad.1348805

Decatur, W.A., Liang, X., Piekna-Przybylska, D., Fournier, M.J., 2007. Identifying effects of snoRNA-guided modifications on the synthesis and function of the yeast ribosome. Methods Enzymol. 425, 283–316. doi:10.1016/S0076-6879(07)25013-X

Dembowski, J.A., Kuo, B., Woolford, J.L., 2013. Has1 regulates consecutive maturation and processing steps for assembly of 60S ribosomal subunits. Nucleic Acids Res. 41, 7889–904. doi:10.1093/nar/gkt545

Du, Y.N., Stillman, B., 2002. Yph1p , an ORC-Interacting Protein : Potential Links between Cell Proliferation Control , DNA Replication , and Ribosome Biogenesis 109, 835–848.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. 66, 486–501. doi:10.1107/S0907444910007493

Fatica, A., Cronshaw, A.D., Tollervey, D., Arbor, A., 2002. Ssf1p Prevents Premature Processing of an Early Pre-60S Ribosomal Particle 9, 341–351.

Fernández-Pevida, A., Kressler, D., de la Cruz, J., 2015. Processing of preribosomal RNA in Saccharomyces cerevisiae. Wiley Interdiscip. Rev. RNA 6, 191–209. doi:10.1002/wrna.1267

Gamalinda, M., Ohmayer, U., Jakovljevic, J., Kumcuoglu, B., Woolford, J., Mbom, B., Lin, L., Jr, J.L.W., 2014. A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains 198–210. doi:10.1101/gad.228825.113.3

Ganley, A.R.D., Kobayashi, T., 2007. Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res. 17, 184–91. doi:10.1101/gr.5457707

Garcia-Higuera, I., Gaitatzes, C., Smith, T.F., Neer, E.J., 1998. Folding a WD repeat propeller. Role of highly conserved aspartic acid residues in the G protein beta subunit and Sec13. J. Biol. Chem. 273, 9041–9.

Gelly, J.-C., Joseph, A.P., Srinivasan, N., de Brevern, A.G., 2011. iPBA: a tool for protein structure comparison using sequence alignment strategies. Nucleic Acids Res. 39, W18–W23. doi:10.1093/nar/gkr333

Gerloff, D.L., Woods, N.T., Farago, A.A., Monteiro, A.N.A., 2012. BRCT domains: A little more than kin, and less than kind. FEBS Lett. 586, 2711–2716. doi:10.1016/j.febslet.2012.05.005

Granneman, S., Lin, C., Champion, E. a, Nandineni, M.R., Zorca, C., Baserga, S.J., 2006. The nucleolar protein Esf2 interacts directly with the DExD/H box RNA helicase, Dbp8, to stimulate ATP hydrolysis. Nucleic Acids Res. 34, 3189–99. doi:10.1093/nar/gkl419

Granneman, S., Petfalski, E., Tollervey, D., 2011. A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. EMBO J. 30, 4006–4019. doi:10.1038/emboj.2011.256

Grimm, T., Hölzel, M., Rohrmoser, M., Harasim, T., Malamoussi, A., Gruber-Eber, A., Kremmer, E., Eick, D., 2006. Dominant-negative Pes1 mutants inhibit ribosomal RNA processing and cell proliferation via incorporation into the PeBoW-complex. Nucleic Acids Res. 34, 3030–43. doi:10.1093/nar/gkl378

Haque, J., Boger, S., Li, J., Duncan, S. a, 2000. The murine Pes1 gene encodes a nuclear protein containing a BRCT domain. Genomics 70, 201–10. doi:10.1006/geno.2000.6375

Harnpicharnchai, P., Jakovljevic, J., Horsey, E., Miles, T., Roman, J., Rout, M., Meagher, D., Imai, B., Guo, Y., Brame, C.J., Shabanowitz, J., Hunt, D.F., Woolford, J.L., 2001. Composition and functional characterization of yeast 66S ribosome assembly intermediates. Mol. Cell 8, 505–15.

Henras, a K., Soudet, J., Gérus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, a, Henry, Y., 2008. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell. Mol. Life Sci. 65, 2334–59. doi:10.1007/s00018-008-8027-0

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.-F., Chakraborty, A., Gleizes, P.-E., 2014. An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip. Rev. RNA. doi:10.1002/wrna.1269

Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S., Tollervey, D., 1994. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. EMBO J. 13, 2452–63.

Holm, L., Rosenström, P., 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–9. doi:10.1093/nar/gkq366

Hölzel, M., Grimm, T., Rohrmoser, M., Malamoussi, A., Harasim, T., Gruber-Eber, A., Kremmer, E., Eick, D., 2007. The BRCT domain of mammalian Pes1 is crucial for nucleolar localization and rRNA processing. Nucleic Acids Res. 35, 789–800. doi:10.1093/nar/gkl1058

Hölzel, M., Rohrmoser, M., Schlee, M., Grimm, T., Harasim, T., Malamoussi, A., Gruber-Eber, A., Kremmer, E., Hiddemann, W., Bornkamm, G.W., Eick, D., 2005. Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. J. Cell Biol. 170, 367–78. doi:10.1083/jcb.200501141

Horiguchi, G., Mollá-Morales, A., Pérez-Pérez, J.M., Kojima, K., Robles, P., Ponce, M.R., Micol, J.L., Tsukaya, H., 2011. Differential contributions of ribosomal protein genes to Arabidopsis thaliana leaf development. Plant J. 65, 724–36. doi:10.1111/j.1365-313X.2010.04457.x

Horn, D.M., Mason, S.L., Karbstein, K., 2011. Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production. J. Biol. Chem. 286, 34082–7. doi:10.1074/jbc.M111.268649

Horsey, E.W., Jakovljevic, J., Miles, T.D., Harnpicharnchai, P., Woolford, J.L., 2004. Role of the yeast Rrp1 protein in the dynamics of pre-ribosome maturation. RNA 10, 813–27. doi:10.1261/rna.5255804

Houtman, J.C.D., Brown, P.H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L.E., Schuck, P., 2007. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. Protein Sci. 16, 30–42. doi:10.1110/ps.062558507

Jakovljevic, J., Ohmayer, U.L.I., Gamalinda, M., Talkish, J., Alexander, L., Linnemann, J.A.N., Milkereit, P., Jr, J.L.W., 2012. Ribosomal proteins L7 and L8 function in concert with six A 3 assembly factors to propagate assembly of domains I and II of 25S rRNA in yeast 60S ribosomal subunits 1805–1822. doi:10.1261/rna.032540.112.40S

Jenner, L., Demeshkina, N., Yusupova, G., Yusupov, M., 2010. Structural rearrangements of the ribosome at the tRNA proofreading step. Nat. Struct. Mol. Biol. 17, 1072–8. doi:10.1038/nsmb.1880

Jenner, L., Starosta, A.L., Terry, D.S., Mikolajka, A., Filonava, L., Yusupov, M., Blanchard, S.C., Wilson, D.N., Yusupova, G., 2013. Structural basis for potent inhibitory activity of the antibiotic tigecycline during protein synthesis. Proc. Natl. Acad. Sci. U. S. A. 110, 3812–6. doi:10.1073/pnas.1216691110

Johnson, A.W., Lund, E., Dahlberg, J., 2002. Nuclear export of ribosomal subunits. Trends Biochem. Sci. 27, 580–585. doi:10.1016/S0968-0004(02)02208-9

Kabsch, W., 2010. Xds. Acta Crystallogr. D. Biol. Crystallogr. 66, 125–32. doi:10.1107/S0907444909047337

Karbstein, K., 2010. Chaperoning ribosome assembly. J. Cell Biol. 189, 11–2. doi:10.1083/jcb.201002102

Kaschner, L.A., Sharma, R., Shrestha, O.K., Meyer, A.E., Craig, E.A., 2015. A conserved domain important for association of eukaryotic J-protein co-chaperones Jjj1 and Zuo1 with the ribosome. Biochim. Biophys. Acta 1853, 1035–45. doi:10.1016/j.bbamcr.2015.01.014

Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C.A., Schuck, P., 2012. High-precision isothermal titration calorimetry with automated peak-shape analysis. Anal. Chem. 84, 5066–73. doi:10.1021/ac3007522

Kellner, M., Rohrmoser, M., Forné, I., Voss, K., Burger, K., Mühl, B., Gruber-Eber, A., Kremmer, E., Imhof, A., Eick, D., 2015. DEAD-box helicase DDX27 regulates 3' end formation of ribosomal 47S RNA and stably associates with the PeBoW-complex. Exp. Cell Res. doi:10.1016/j.yexcr.2015.03.017

Killian, A., Le Meur, N., Sesboüé, R., Bourguignon, J., Bougeard, G., Gautherot, J., Bastard, C., Frébourg, T., Flaman, J.-M., 2004. Inactivation of the RRB1-Pescadillo pathway involved in ribosome biogenesis induces chromosomal instability. Oncogene 23, 8597–602. doi:10.1038/sj.onc.1207845

Kinoshita, Y., Jarell, a D., Flaman, J.M., Foltz, G., Schuster, J., Sopher, B.L., Irvin, D.K., Kanning, K., Kornblum, H.I., Nelson, P.S., Hieter, P., Morrison, R.S., 2001. Pescadillo, a novel cell cycle regulatory protein abnormally expressed in malignant cells. J. Biol. Chem. 276, 6656–65. doi:10.1074/jbc.M008536200

Korobeinikova, A. V, Garber, M.B., Gongadze, G.M., 2012. Ribosomal Proteins : Structure , Function , and Evolution 77, 562–574.

Kos, M., Tollervey, D., 2010. Yeast pre-rRNA processing and modification occur cotranscriptionally. Mol. Cell 37, 809–20. doi:10.1016/j.molcel.2010.02.024

Kressler, D., Hurt, E., Bassler, J., 2010. Driving ribosome assembly. Biochim. Biophys. Acta 1803, 673–83. doi:10.1016/j.bbamcr.2009.10.009

Kressler, D., Hurt, E., Bergler, H., Bassler, J., 2012. The power of AAA-ATPases on the road of pre-60S ribosome maturation--molecular machines that strip pre-ribosomal particles. Biochim. Biophys. Acta 1823, 92–100. doi:10.1016/j.bbamcr.2011.06.017

Kressler, D., Roser, D., Pertschy, B., Hurt, E., 2008. The AAA ATPase Rix7 powers progression of ribosome biogenesis by stripping Nsa1 from pre-60S particles. J. Cell Biol. 181, 935–44. doi:10.1083/jcb.200801181

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–97. doi:10.1016/j.jmb.2007.05.022

Kubota, S., Kubota, H., Nagata, K., 2006. Cytosolic chaperonin protects folding intermediates of Gbeta from aggregation by recognizing hydrophobic beta-strands. Proc. Natl. Acad. Sci. U. S. A. 103, 8360–8365. doi:10.1073/pnas.0600195103

Lafontaine, D.L.J., 2015. Noncoding RNAs in eukaryotic ribosome biogenesis and function. Nat. Struct. Mol. Biol. 22, 11–19. doi:10.1038/nsmb.2939

Lamanna, A.C., Karbstein, K., 2009. Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. Proc. Natl. Acad. Sci. U. S. A. 106, 14259–64. doi:10.1073/pnas.0905403106

Lapik, Y.R., Fernandes, C.J., Lau, L.F., Pestov, D.G., 2004. Physical and Functional Interaction between Pes1 and Bop1 in Mammalian Ribosome Biogenesis University of Illinois at Chicago 15, 17–29.

Leslie, A.W., Powell, H., 2007. Processing diffraction data with mosflm, in: Read, R., Sussman, J. (Eds.), Evolving Methods for Macromolecular Crystallography SE - 4, NATO Science Series. Springer Netherlands, pp. 41–51. doi:10.1007/978-1-4020-6316-9_4

Long, F., Vagin, A. a, Young, P., Murshudov, G.N., 2008. BALBES: a molecular-replacement pipeline. Acta Crystallogr. D. Biol. Crystallogr. 64, 125–32. doi:10.1107/S0907444907050172

Matsuo, Y., Granneman, S., Thoms, M., Manikas, R.-G., Tollervey, D., Hurt, E., 2014. Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. Nature 505, 112–6. doi:10.1038/nature12731

Mccann, K.L., Charette, J.M., Vincent, N.G., Baserga, S.J., 2015. A protein interaction map of the LSU processome. Genes Dev. 29, 862–75. doi:10.1101/gad.256370.114

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., Yusupov, M., 2012. One core, two shells: bacterial and eukaryotic ribosomes. Nat. Struct. Mol. Biol. 19, 560–7. doi:10.1038/nsmb.2313

Meng, E.C., Pettersen, E.F., Couch, G.S., Huang, C.C., Ferrin, T.E., 2006. Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinformatics 7, 339. doi:10.1186/1471-2105-7-339

Merl, J., Jakob, S., Ridinger, K., Hierlmeier, T., Deutzmann, R., Milkereit, P., Tschochner, H., 2010. Analysis of ribosome biogenesis factor-modules in yeast cells depleted from pre-ribosomes. Nucleic Acids Res. 38, 3068–80. doi:10.1093/nar/gkp1244

Meyer, A.E., Hung, N.-J., Yang, P., Johnson, A.W., Craig, E.A., 2007. The specialized cytosolic J-protein, Jjj1, functions in 60S ribosomal subunit biogenesis. Proc. Natl. Acad. Sci. U. S. A. 104, 1558–1563. doi:10.1073/pnas.0610704104

Miles, T.D., Jakovljevic, J., Horsey, E.W., Harnpicharnchai, P., Tang, L., Woolford, J.L., 2005. Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes. Mol. Cell. Biol. 25, 10419–32. doi:10.1128/MCB.25.23.10419-10432.2005

Moss, T., Langlois, F., Gagnon-Kugler, T., Stefanovsky, V., 2007. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell. Mol. Life Sci. 64, 29–49. doi:10.1007/s00018-006-6278-1

Mushtaq, A., Jamill, A., 2012. Cloning of a β -glucosidase gene from thermophilic fungus Cheatomium thermophilum. Pakistan J. Life Soc. Sci. 10, 98–101.

Mylona, A., Fernández-Tornero, C., Legrand, P., Haupt, M., Sentenac, A., Acker, J., Müller, C.W., 2006. Structure of the tau60/Delta tau91 subcomplex of yeast transcription factor IIIC: insights into preinitiation complex assembly. Mol. Cell 24, 221–32. doi:10.1016/j.molcel.2006.08.013

Nal, B., Mohr, E., Silva, M.-I. Da, Tagett, R., Navarro, C., Carroll, P., Depetris, D., Verthuy, C., Jordan, B.R., Ferrier, P., 2002. Wdr12, a mouse gene encoding a novel WD-Repeat Protein with a notchless-like amino-terminal domain. Genomics 79, 77–86. doi:10.1006/geno.2001.6682

Nissan, T.A., Baûler, J., Petfalski, E., Tollervey, D., Hurt, E., 2002. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm 21, 5539–5547.

Notredame, C., Higgins, D.G., Heringa, J., 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302, 205–17. doi:10.1006/jmbi.2000.4042

Oeffinger, M., Lueng, A., Lamond, A., Tollervey, D., 2002. Yeast Pescadillo is required for multiple activities during 60S ribosomal subunit synthesis. Rna 8, 626–636. doi:10.1017/S1355838202020022

Oeffinger, M., Zenklusen, D., Ferguson, A., Wei, K.E., El Hage, A., Tollervey, D., Chait, B.T., Singer, R.H., Rout, M.P., 2009. Rrp17p is a eukaryotic exonuclease required for 5' end processing of Pre-60S ribosomal RNA. Mol. Cell 36, 768–81. doi:10.1016/j.molcel.2009.11.011

Ohmayer, U., Gamalinda, M., Sauert, M., Ossowski, J., Pöll, G., Linnemann, J., Hierlmeier, T., Perez-Fernandez, J., Kumcuoglu, B., Leger-Silvestre, I., Faubladier, M., Griesenbeck, J., Woolford, J., Tschochner, H., Milkereit, P., 2013. Studies on the assembly characteristics of large subunit ribosomal proteins in S. cerevisae. PLoS One 8, e68412. doi:10.1371/journal.pone.0068412

Ouspenski, I.I., Elledge, S.J., Brinkley, B.R., 1999. New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability 27, 3001–3008.

Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P., Salemme, F.R., 2001. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J. Biomol. Screen. 6, 429–40. doi:10.1089/108705701753364922

Parcellier, A., Schmitt, E., Gurbuxani, S., Seigneurin-Berny, D., Pance, A., Chantôme, A., Plenchette, S., Khochbin, S., Solary, E., Garrido, C., 2003. HSP27 is a ubiquitin-binding protein involved in I-kappaBalpha proteasomal degradation. Mol. Cell. Biol. 23, 5790–802. doi:10.1128/MCB.23.16.5790–5802.2003

Pestov, D.G., Stockelman, M.G., Strezoska, Z., Lau, L.F., 2001a. ERB1, the yeast homolog of mammalian Bop1, is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs. Nucleic Acids Res. 29, 3621–30.

Pestov, D.G., Strezoska, Z., Lau, L.F., 2001b. Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: effects of nucleolar protein Bop1 on G(1)/S transition. Mol. Cell. Biol. 21, 4246–55. doi:10.1128/MCB.21.13.4246-4255.2001

Phillips, A.H., Zhang, Y., Cunningham, C.N., Zhou, L., Forrest, W.F., Liu, P.S., Steffek, M., Lee, J., Tam, C., Helgason, E., Murray, J.M., Kirkpatrick, D.S., Fairbrother, W.J., Corn, J.E., 2013. Conformational dynamics control ubiquitin-deubiquitinase interactions and influence in vivo signaling. Proc. Natl. Acad. Sci. U. S. A. 110, 11379–84. doi:10.1073/pnas.1302407110

Phipps, K.R., Charette, J.M., Baserga, S.J., 2012. The small subunit processome in ribosome biogenesis—progress and prospects. Wiley Interdiscip. Rev. RNA 2, 1–21. doi:10.1002/wrna.57

Ramakrishnan, V., 2002. Ribosome structure and the mechanism of translation. Cell 108, 557–72.

Rodríguez-Galán, O., García-Gómez, J.J., de la Cruz, J., 2013. Yeast and human RNA helicases involved in ribosome biogenesis: current status and perspectives. Biochim. Biophys. Acta 1829, 775–90. doi:10.1016/j.bbagrm.2013.01.007

Rodríguez-Mateos, M., Abia, D., García-Gómez, J.J., Morreale, A., de la Cruz, J., Santos, C., Remacha, M., Ballesta, J.P.G., 2009. The amino terminal domain from Mrt4 protein can functionally replace the RNA binding domain of the ribosomal P0 protein. Nucleic Acids Res. 37, 3514–21. doi:10.1093/nar/gkp209

Rohrmoser, M., Ho, M., Grimm, T., Malamoussi, A., Harasim, T., Orban, M., Pfisterer, I., Gruber-Eber, A., Kremmer, E., Eick, D., Hölzel, M., 2007. Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. Mol. Cell. Biol. 27, 3682–94. doi:10.1128/MCB.00172-07

Rosado, I. V, de la Cruz, J., 2004. Npa1p is an essential trans-acting factor required for an early step in the assembly of 60S ribosomal subunits in Saccharomyces cerevisiae. RNA 10, 1073–1083. doi:10.1261/rna.7340404

Russell, J., Zomerdijk, J.C.B.M., 2006. The RNA polymerase I transcription machinery. Biochem. Soc. Symp. 203–16.

Sahasranaman, A., Dembowski, J., Strahler, J., Andrews, P., Maddock, J., Woolford, J.L., 2011. Assembly of Saccharomyces cerevisiae 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing. EMBO J. 30, 4020–32. doi:10.1038/emboj.2011.338

Sakumoto, N., Yamashita, H., Mukai, Y., Kaneko, Y., Harashima, S., 2001. Dual-specificity protein phosphatase Yvh1p, which is required for vegetative growth and sporulation, interacts with yeast pescadillo homolog in Saccharomyces cerevisiae. Biochem. Biophys. Res. Commun. 289, 608–15. doi:10.1006/bbrc.2001.6021

Saveanu, C., Namane, A., Gleizes, P., Lebreton, A., Rousselle, J., Noaillac-depeyre, J., Gas, N., Jacquier, A., Fromont-racine, M., 2003. Sequential Protein Association with Nascent 60S Ribosomal Particles 23, 4449–4460. doi:10.1128/MCB.23.13.4449

Saveanu, C., Rousselle, J.-C., Lenormand, P., Namane, A., Jacquier, A., Fromont-Racine, M., 2007. The p21-activated protein kinase inhibitor Skb15 and its budding yeast homologue are 60S ribosome assembly factors. Mol. Cell. Biol. 27, 2897–909. doi:10.1128/MCB.00064-07

Schmeing, T.M., Ramakrishnan, V., 2009. What recent ribosome structures have revealed about the mechanism of translation. Nature 461, 1234–42. doi:10.1038/nature08403

Schuldt, A., 2011. Gene expression: personalized ribosomes. Nat. Rev. Mol. Cell Biol. 12, 344–5. doi:10.1038/nrm3127

Shen, J., Cowen, L.E., Griffin, A.M., Chan, L., Ko, J.R., 2008. The Candida albicans pescadillo homolog is required for normal hypha-to-yeast morphogenesis and.

Smith, T.F., Gaitatzes, C., Saxena, K., Neer, E.J., 1999. The WD repeat: a common architecture for diverse functions. Trends Biochem. Sci. 24, 181–5.

Sørensen, P.D., Frederiksen, S., 1991. Characterization of human 5S rRNA genes. Nucleic Acids Res. 19, 4147–4151.

Steffen, K.K., McCormick, M.A., Pham, K.M., MacKay, V.L., Delaney, J.R., Murakami, C.J., Kaeberlein, M., Kennedy, B.K., 2012. Ribosome deficiency protects against ER stress in Saccharomyces cerevisiae. Genetics 191, 107–18. doi:10.1534/genetics.111.136549

Steitz, T. a, 2008. A structural understanding of the dynamic ribosome machine. Nat. Rev. Mol. Cell Biol. 9, 242–53. doi:10.1038/nrm2352

Stirnimann, C.U., Petsalaki, E., Russell, R.B., Müller, C.W., 2010. WD40 proteins propel cellular networks. Trends Biochem. Sci. 35, 565–74. doi:10.1016/j.tibs.2010.04.003

Strezoska, Z., Pestov, D.G., Lau, L.F., 2000. Bop1 is a mouse WD40 repeat nucleolar protein involved in 28S and 5. 8S RRNA processing and 60S ribosome biogenesis. Mol. Cell. Biol. 20, 5516–28.

Strezoska, Z., Pestov, D.G., Lau, L.F., 2002. Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. J. Biol. Chem. 277, 29617–25. doi:10.1074/jbc.M204381200

Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. 41, 207–234. doi:10.1016/j.pep.2005.01.016

Talkish, J., Zhang, J., Jakovljevic, J., Horsey, E.W., Woolford, J.L., 2012. Hierarchical recruitment into nascent ribosomes of assembly factors required for 27SB pre-rRNA processing in Saccharomyces cerevisiae. Nucleic Acids Res. 40, 8646–61. doi:10.1093/nar/gks609

Tang, L., Sahasranaman, A., Jakovljevic, J., Schleifman, E., Woolford, J.L., 2008. Interactions among Ytm1, Erb1, and Nop7 required for assembly of the Nop7-subcomplex in yeast preribosomes. Mol. Biol. Cell 19, 2844–56. doi:10.1091/mbc.E07-12-1281

Tecza, A., Bugner, V., Kühl, M., Kühl, S.J., 2011. Pescadillo homologue 1 and Peter Pan function during Xenopus laevis pronephros development. Biol. Cell 103, 483–98. doi:10.1042/BC20110032

Thomson, E., Tollervey, D., 2010. The final step in 5.8S rRNA processing is cytoplasmic in Saccharomyces cerevisiae. Mol. Cell. Biol. 30, 976–84. doi:10.1128/MCB.01359-09

Turowski, T.W., Lebaron, S., Zhang, E., Peil, L., Dudnakova, T., Petfalski, E., Granneman, S., Rappsilber, J., Tollervey, D., 2014. Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. Nucleic Acids Res. 42, 12189–99. doi:10.1093/nar/gku878

Ulbrich, C., Diepholz, M., Bassler, J., Kressler, D., Pertschy, B., Galani, K., Böttcher, B., Hurt, E., Baßler, J., Bo, B., 2009. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. Cell 138, 911–22. doi:10.1016/j.cell.2009.06.045

Venema, J., Planta, R., Raué, H., 1998. In Vivo Mutational Analysis of Ribosomal RNA in Saccharomyces cerevisiae, in: Martin, R. (Ed.), Protein Synthesis SE - 19, Methods in Molecular Biology. Springer New York, pp. 257–270. doi:10.1385/0-89603-397-X:257

Wang, Y., Jiang, F., Zhuo, Z., Wu, X.-H., Wu, Y.-D., 2013. A method for WD40 repeat detection and secondary structure prediction. PLoS One 8, e65705. doi:10.1371/journal.pone.0065705

Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. Trends Biochem. Sci. 24, 437–440. doi:10.1016/S0968-0004(99)01460-7

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., Barton, G.J., 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–91. doi:10.1093/bioinformatics/btp033

Willmund, F., del Alamo, M., Pechmann, S., Chen, T., Albanèse, V., Dammer, E.B., Peng, J., Frydman, J., 2013. The cotranslational function of ribosome-associated Hsp70 in eukaryotic protein homeostasis. Cell 152, 196–209. doi:10.1016/j.cell.2012.12.001

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., Ramakrishnan, V., 2000. Structure of the 30S ribosomal subunit. Nature 407, 327–39. doi:10.1038/35030006

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E. a, Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. 67, 235–42. doi:10.1107/S0907444910045749

Winter, G., 2009. xia2 : an expert system for macromolecular crystallography data reduction. J. Appl. Crystallogr. 43, 186–190. doi:10.1107/S0021889809045701

Wong, Q.W.-L., Li, J., Ng, S.R., Lim, S.G., Yang, H., Vardy, L.A., 2014. RPL39L is an example of a recently evolved ribosomal protein paralog that shows highly specific tissue expression patterns and is upregulated in ESCs and HCC tumors. RNA Biol. 11, 33–41. doi:10.4161/rna.27427

Woolford, J.L., Baserga, S.J., 2013. Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics 195, 643–81. doi:10.1534/genetics.113.153197

Wu, X.-H., Wang, Y., Zhuo, Z., Jiang, F., Wu, Y.-D., 2012. Identifying the hotspots on the top faces of WD40-repeat proteins from their primary sequences by β-bulges and DHSW tetrads. PLoS One 7, e43005. doi:10.1371/journal.pone.0043005

Xu, C., Min, J., 2011. Structure and function of WD40 domain proteins. Protein Cell 2, 202–14. doi:10.1007/s13238-011-1018-1

Xu, Y., Liu, Y., Lee, J., Ye, Y., 2013. A ubiquitin-like domain recruits an oligomeric chaperone to a retrotranslocation complex in endoplasmic reticulum-associated degradation. J. Biol. Chem. 288, 18068–76. doi:10.1074/jbc.M112.449199

Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R., Barna, M., 2015. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. Nature 517, 33–8. doi:10.1038/nature14010

Yelick, P.C., Trainor, P.A., 2015. Ribosomopathies: Global process, tissue specific defects. Rare Dis. 3, e1025185. doi:10.1080/21675511.2015.1025185

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H., Noller, H.F., 2001. Crystal structure of the ribosome at 5.5 A resolution. Science 292, 883–96. doi:10.1126/science.1060089

Zemp, I., Kutay, U., 2007. Nuclear export and cytoplasmic maturation of ribosomal subunits. FEBS Lett. 581, 2783–93. doi:10.1016/j.febslet.2007.05.013

Zhang, J., Yang, Y., Wu, J., 2009. B23 interacts with PES1 and is involved in nucleolar localization of PES1 41, 991–997. doi:10.1093/abbs/gmp096.Advance

Zhao, Y., Chapman, D.A.G., Jones, I.M., 2003. Improving baculovirus recombination. Nucleic Acids Res. 31, e6–e6.

Zografidis, A., Kapolas, G., Podia, V., Beri, D., Papadopoulou, K., Milioni, D., Haralampidis, K., 2014. Transcriptional regulation and functional involvement of the Arabidopsis pescadillo ortholog AtPES in root development. Plant Sci. 229, 53–65. doi:10.1016/j.plantsci.2014.08.012

8. APPENDIX

8.1. Abbreviations

3C: Human Rhinovirus protease
40S: Small Eukaryotic Ribosomal Subunit (SSU)
6xHis: 6x Histidine tag
60S: Large Eukaryotic Ribosomal Subunit (LSU)
APBS: Adaptive Poisson-Boltzmann Solver
BSA: Bovine serum albumin
CCP4: Collaborative Computational Project Number 4
ChErb1: Erb1 protein from *C. thermophilum*
ChErb1Ct: Carboxy-terminal domain of ChErb1 protein from *C. thermophilum*
ChNop7: Nop7 protein from *C.thermophilum*
ChYtm1: Ytm1 protein from *C. thermophilum*
CRAC: UV cross-linking and analysis of cDNA
C-terminal, C-term, Ct: Carboxy terminal
DNA: Deoxyribonucleic acid
dNTP: Deoxyribonucleotide triphosphate
DTT: Dithiothreitol
Erb1Ct: Carboxy-terminal domain of Erb1 protein
fl: Full length
5-FOA: 5-fluoroorotic acid
GST: Glutathione S-transferase
HMM: Hidden Markov Model
IMAC: Immobilized-metal affinity chromatography
IPTG: Isopropyl β-D-1-thiogalactopyranoside
K_D: Dissociation constant
LB: Lysogeny Broth
LSU: Large Eukaryotic Ribosomal Subunit (60S)
pI: Isoelectric point
MBP: Maltose Binding Protein
MR: Molecular replacement
MW: Molecular Weight
Ni-NTA: Nickel-nitriloacetic acid
NMR: Nuclear magnetic resonance
N-terminal, N-term: Amino terminal
OD: Optical density at 600 nm
PCR: Polymerase chain reaction
PDB: Protein data bank
PEST: proline, glutamic acid, serine and threonine rich sequence

PolyU: Poly uridylic acid
RMSD: Root mean square deviation
RNA: Ribonucleic acid
S: Svedberg unit (sedimentation)
SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSU: Small Eukaryotic Ribosomal Subunit (40S)
SUMO: Small ubiquitin-like modifier
TEV: Tobacco Etch Virus protease
TLS: Translation-libration-screw
Tm: Melting temperature
UBL: Ubiquitin-like Domain
UV: Ultra violet
WD40: 40-residue long repeat containing tryptophan-aspartic acid dipeptide
YPD: yeast extract, peptone, dextrose containing medium

8.2. Publications

PLOS | ONE

RESEARCH ARTICLE

The Carboxy-Terminal Domain of Erb1 Is a Seven-Bladed β-Propeller that Binds RNA

Wegrecki Marcin[1], Jose Luis Neira[2,3], Jeronimo Bravo[1]*

1 Instituto de Biomedicina de Valencia, Consejo Superior de Investigaciones Científicas, c/ Jaime Roig 11, 46010 Valencia, Spain, 2 Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202 Elche (Alicante), Spain, 3 Instituto de Biocomputación y Física de los Sistemas Complejos (BIFI), 50009 Zaragoza, Spain

* jbravo@ibv.csic.es

OPEN ACCESS

Citation: Marcin W, Neira JL, Bravo J (2015) The Carboxy-Terminal Domain of Erb1 Is a Seven-Bladed ß-Propeller that Binds RNA. PLoS ONE 10(4): e0123463. doi:10.1371/journal.pone.0123463

Academic Editor: Maria Sola, Molecular Biology Institute of Barcelona, CSIC, SPAIN

Received: December 2, 2014

Accepted: March 4, 2015

Published: April 16, 2015

Data Availability Statement: The atomic coordinates and structure factors have been deposited in the PDB (http://www.rcsb.org/) with the accession number: 4U7A.

Funding: This work was supported by the European Community, Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement N° 283570) to JB; Ministerio de Economía y Competitividad, SAF2012-31405 to JB and CSD2008-00005 and CTQ2013-4493 to JLN; and Generalitat Valenciana PROMETEO/2012/061 (JB) and 2013/018 (JLN). MW received a JAE-PREDOC fellowship from Consejo Superior de Investigaciones Científicas, Spain. The funders had no role in study

Abstract

Erb1 (Eukaryotic Ribosome Biogenesis 1) protein is essential for the maturation of the ribosomal 60S subunit. Functional studies in yeast and mammalian cells showed that altogether with Nop7 and Ytm1 it forms a stable subcomplex called PeBoW that is crucial for a correct rRNA processing. The exact function of the protein within the process remains unknown. The N-terminal region of the protein includes a well conserved region shown to be involved in PeBoW complex formation whereas the carboxy-terminal half was predicted to contain seven WD40 repeats. This first structural report on Erb1 from yeast describes the architecture of a seven-bladed β-propeller domain that revealed a characteristic extra motif formed by two α-helices and a β-strand that insert within the second WD repeat. We performed analysis of molecular surface and crystal packing, together with multiple sequence alignment and comparison of the structure with other β-propellers, in order to identify areas that are more likely to mediate protein-protein interactions. The abundance of many positively charged residues on the surface of the domain led us to investigate whether the propeller of Erb1 might be involved in RNA binding. Three independent assays confirmed that the protein interacted in vitro with polyuridilic acid (polyU), thus suggesting a possible role of the domain in rRNA rearrangement during ribosome biogenesis.

Introduction

Erb1/Bop1 is a eukaryotic protein that was firstly described as an evolutionary conserved factor involved in large ribosomal subunit biogenesis in yeast and mammals respectively [1,2]. Its function is essential in the processing of rRNA precursors that give rise to the mature 5.8S and 25/28S particles [1–3]. Knock-down of Erb1 impairs ribosome assembly, leading to accumulation of immature rRNA species in yeast, whereas the overexpression of Bop1 negatively affects cell proliferation in mammals [1,3,4]. In addition, an N-terminally truncated mutant of Bop1 is able to induce a reversible growth arrest through p53 response, suggesting a possible role of the protein in ribosome biogenesis control [1,3]. Moreover, over-expression of *bop1* increases the number of multipolar spindles, implying a correlation with colorectal cancer [5].

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

In *Saccharomyces cerevisiae* Erb1 contains 807 residues and carries a well conserved N-terminal domain called BOP1NT which plays role in the recruitment of the protein to pre-ribosomes [6]. The C-terminal region of Erb1 was predicted to contain seven WD repeats that form a β-propeller domain of unclear function [1]. Additional work on the exact role of Erb1 in ribosome assembly showed that it formed part of a functional cluster of processing factors, called A3, that were responsible for the cleavage of ITS1 (Internal Transcribed Spacer 1) [7,8]. It has been also demonstrated that the full length protein binds to Domain I of 25S rRNA [9]. Erb1 directly interacts with Nop7 and Ytm1 proteins (Pes1 and Wdr12 in mammals, respectively) forming Nop7 sub-complex (called PeBoW in mammals) that co-purifies with pre60S particles but remains stable even after its dissociation from pre-ribosomes [7,10,11]. Nop7 complex has to be removed from the nascent ribosome by the AAA-ATPase Rea1 in order to promote normal ribosome maturation [12]. Since Ytm1 and Nop7 do not physically interact, Erb1 is considered to be the core of the complex and the ratio of Nop7/Erb1 and Erb1/Ytm1 heterodimers is important in controlling the assembly and function of Nop7 complex (as shown for PeBoW complex in mammals by Rohrmoser [4]). The involvement of the complex in ribosome biogenesis was reviewed by Henras [13].

While several studies regarding Erb1 function and interactions focus on the BOP1NT domain, the role of the propeller is still under investigation. It was shown that in yeast a truncated Erb1 lacking the C-terminal domain would not cause growth arrest but presented only a mild defect in rRNA processing [6]. Despite the fact that the β-propeller domain of Erb1 has been proposed as dispensable for ribosome assembly, it still presents a high degree of conservation in all eukaryotes. It is worth noting that the binding partner of Erb1, Ytm1 is also predicted to contain a large 7-bladed β-propeller region on its C-terminus [14]. Furthermore, there are described additional 20 proteins that contain β-propeller domains in their structures and form part of the ribosome assembly pathway in eukaryotes, thus indicating that it is a common fold required to establish a high-affinity protein-protein interaction network within this complex pathway [13].

In recent years, there has been an increasing interest in the architecture of pre-ribosomes in order to get a better understanding of the dynamics of the process. However, there is very limited amount of information regarding the pre-ribosomal particles from a structural point of view. The main challenge in the field is the lack of stability of the individual components of this enormous machinery as well as the difficulty when trying to obtain homogenous samples for structural studies [15]. Thanks to the recent advances in cryo-EM technique it has been possible to get an insight into the organization of the late-stage pre-ribosome, nevertheless the structure of the majority of the factors that participate in ribosome maturation still remains unknown [16,17].

Here we present the structure of the β-propeller domain of Erb1 at 1.6Å resolution that was obtained during crystallization trials of Erb1/Nop7 dimer. The structural information allows us to exactly define the boundaries of the domain and to describe its particular features, being the presence of a long insertion within the second WD repeat the most distinctive characteristic. We consider a possible role of this extra fold in protein-protein interactions based on its importance for crystal packing. At the same time, surface analysis helps us to predict other areas that are likely to be involved in recognition of proteins or nucleic acids therefore making the C-terminal domain of Erb1 a motif capable of binding to additional factors within pre60S network. At last, we demonstrate that, indeed, the β-propeller of Erb1 is able to bind non-specifically RNA in vitro through a saturable surface.

Materials and Methods

Cloning

Since the genes of *nop7* and *erb1* do not contain introns, they were cloned from the genomic DNA of *Saccharomyces cerevisiae*. A second PCR was performed in order to add overhangs

suitable for ligase independent cloning (LIC) cloning and compatible with vectors of interest. *Nop7* was cloned into pNIC28-Bsa4 vector which contains a sequence for N-terminal 6xHis tag followed by TEV (Tobacco Etch Virus protease) cleavage site. *Erb1* was introduced into pET28-NKI/LIC 6His/3C vector obtained from Dr A. Perrakis group (NKI, Amsterdam), containing the N-terminal 6xHis tag followed by 3C protease (Human Rhinovirus protease) cleavage site. Both ligation reactions were performed according to the standard LIC protocol using T4 DNA polymerase from Fermentas. The DNA coding for yeast $Erb1_{518\text{-}586}$ containing LIC suitable overhangs was purchased from Life Technologies and the gene of $Erb1_{432\text{-}801}$ from *Chaetomium thermophilum var thermphilum* (ChErb1) was cloned from the cDNA library prepared as described in [18]. Both genes were cloned into pET28-NKI/LIC 6His/3C as described above.

Protein expression and purification

The 6xHis-Nop7 was expressed in BL21 Codon Plus strain of *E. coli* grown in LB. The expression was induced with 0.5mM IPTG and the culture was incubated overnight at 20°C. *E. coli* BL21 Codon Plus strain was also used for the expression of 6xHis-Erb1. Protein production was performed overnight at 20°C using ZY medium and autoinduction strategy described by Studier [19]. After 16h of expression the cells were collected by centrifugation and the resulting pellets were frozen in liquid nitrogen and stored at -80°C. Both proteins were purified separately following a protocol that involved IMAC (Immobilized-Metal Affinity Chromatography) and size exclusion chromatography. The cell pellet was resuspended in 40ml of lysis buffer containing 50mM Hepes pH 7.5; 0.5M NaCl; 10% glycerol; 5mM β-mercaptoethanol and 10mM of imidazole. One pill of Roche EDTA-free protease inhibitor cocktail was added to the buffer. The cells were disrupted by sonication and the soluble fraction was separated from the debris by centrifugation at 4°C/20000 x g for 45 minutes. The supernatant was filtered and loaded onto a 5ml HisTrap column (GE Healthcare) previously equilibrated with 10 volumes of Wash Buffer (20mM Hepes pH 7.5; 0.5M NaCl; 2mM β-mercaptoethanol; 5% glycerol and 20mM imidazole). The wash and elution steps were performed using the AKTA purification system (GE Healthcare) by following absorbance at 280 nm. First, the column was washed with 10 column volumes of Wash Buffer and then the protein was eluted in 5ml fractions using a step gradient of 3%, 12%, 40% and 100% of elution Buffer (20mM Hepes pH 7.5; 0.5M NaCl; 2mM β-mercaptoethanol; 5% glycerol and 500mM imidazole). The fractions containing the protein of interest were concentrated and injected into a Superdex200 16/60 column which was equilibrated with 1.5 column volume of SE buffer (20mM Hepes pH 7.5; 0.25M NaCl; 2 mM β-mercaptoethanol, 5% glycerol). The 5ml fractions were collected and those containing non-aggregated protein were concentrated.

Expression and purification of $ChErb1_{432\text{-}801}$ were performed as described for Erb1from *S. cerevisiae* except for the size exclusion chromatography step. The buffer used for $ChErb1_{432\text{-}801}$ was 50mM Tris pH8; 100mM NaCl and 5mM $MgCl_2$. The fractions containing the protein were concentrated up to 4mg/ml.

Protein expression and purification of $Erb1_{518\text{-}586}$

The 6xHis-NKI was expressed in C41 (Lucigen, USA) strain of *E. coli* grown in 5l of LB. The expression was induced with 1mM IPTG, when the dispersion at 600 nm reached a value of 0.6–0.8, and the culture was incubated overnight at 37°C. After 16h of expression the cells were collected by centrifugation and the resulting pellets were frozen in liquid nitrogen and stored at -80°C. The cell pellet was resuspended in 50ml of lysis buffer (20 mM Tris pH 8.0; 0.5 M NaCl; 1mM β-mercaptoethanol; 0.1% Triton X-100 and 5mM of imidazole), supplemented with one tablet of Sigma protease inhibitor cocktail (EDTA-free). The cells were disrupted by sonication

(10 burst of 45 s at the maximum power of the sonicator, Model 102-C, Branson, USA), in ice, and the soluble fraction was separated from the debris by centrifugation at 4°C/20000g for 45 minutes. The supernatant was filtered and added to a 5 ml of His-select nickel affinity gel (Sigma). The resulting mixture was incubated for 30 minutes at 4°C; after that time it was added to a Biorad empty column and the supernatant was separated by gravity. The resin was washed with 20 ml of 20 mM Tris pH 8.0; 0.5 M NaCl; 1mM β-mercaptoethanol and 25 mM of imidazole for 10 minutes. The supernatant was removed from the column by gravity. The protein was eluted from the resin with 20 ml of 20 mM Tris pH 8.0; 0.5 M NaCl; 1mM β-mercaptoethanol and 500 mM of imidazole. The presence of the protein in the eluate was confirmed by 18% SDS-PAGE gels. The protein was concentrated in Amicon centrifugal devices (Amicon, MW cutoff 3000 Da) and loaded into a gel filtration column Superdex 75 16/600 (GE Healthcare) coupled to a FPLC purification system (GE Healthcare), by following the absorbance at 280 nm. The column was equilibrated in buffer 50 mM Tris (pH 7.5) with 150 mM NaCl. For ^{15}N-labelled samples, the M9 minimal medium was used to express the protein. Purification protocols were carried out as in LB medium.

Complex formation and co-crystallization

After the size exclusion chromatography step, concentrated samples of Nop7 and Erb1 were mixed in equimolar amounts and injected into Superdex200 16/60 column equilibrated with SE Buffer. The 5ml fractions corresponding to the heterodimer were mixed, concentrated and used for crystallization. Initial crystallization trials were performed at 21°C, the concentration of the Nop7/Erb1 complex was 80mg/ml and drops containing 0.3μl of protein sample and 0.3μl of reservoir were set up. Crystals diffracting up to 2.9Å were obtained in 0.1M Hepes pH 7.5; 10% Polyethylene glycol8000 and 8% ethylene glycol. In order to improve crystal size and resolution we performed an optimization screen based on the Hampton Additives kit, the protein concentration used was 60mg/ml and the drop size of 0.5μl for protein sample, 0.5μl for reservoir and 0.1μl of additive was added.

Data collection and processing

Crystals obtained in the original screening and from the Hampton Additives screening were flash-cooled in liquid nitrogen and diffracted at Diamond Light Source (Harwell, UK) I24 and I03 beamlines. The maximum diffraction up to 1.6Å was obtained with crystals grown with addition of 30% of ethanol The crystals contained one molecule per asymmetric unit and the space group was P 21 21 21 with the following unit cell parameters: a = 52.026Å, b = 62.432Å, c = 158.22Å; $\alpha = \beta = \gamma = 90.00°$. The data were processed using XDS, merged and scaled in CCP4 [20,21]. Molecular replacement was done in parallel using MR module of Phenix and Balbes on-line MR suite [22,23]. The search model used in phenix.mr was a poly-Ala β-propeller based on the input model chosen by Balbes database search (PDB: 2H13), which corresponds to an unrelated WD40 protein, WDR5, engaged in histone binding. Obtained model and initial phases were then used for additional model building cycle by AutoBuild module from Phenix. The final structure was then refined combining phenix.refine suite and manual refining in Coot until the final r factors were R = 16.0% and R free = 17.4%. Data collection and refinement statistics are shown in Table 1. The model and structure factors were deposited in Protein Data Bank with PDB ID: 4U7A.

Circular dichroism

Spectra of Erb1$_{518\text{-}586}$ were collected on a Jasco J810 (Japan) spectropolarimeter connected to a Peltier unit. The instrument was periodically calibrated with (+)-10-camphorsulfonic acid.

Spectra were acquired at 25°C in phosphate buffer at pH 7.0 (10mM). For each experiment, corresponding blank solutions were subtracted. The response time was 2 s, and experiments were averaged over 6 scans, with a scan speed of 50 nm/min. The step resolution was 0.2 nm, and the band-width was 1 nm. Molar ellipticity was obtained as described [24]. We explored a wide range of protein concentrations (10–80 μM) to ascertain whether shape or intensity of the

Table 1. Data collection and refinement statistics.

Wavelength (Å)	0.9763
Resolution range (Å)	58.07–1.6 (1.657–1.6)[a]
Space group	P 21 21 21
Unit cell	
a, b, c (Å)	52.02, 62.43, 158.22
α, β, γ (°)	90, 90, 90
Total reflections	446057 (43775)
Unique reflections	67582 (6545)
Multiplicity	6.6 (6.7)
Completeness (%)	98.07 (96.78)
Mean I/sigma(I)	19.48 (2.16)[b]
Wilson B-factor	22.85
R-merge	0.04807 (0.8424)[c]
R-meas	0.0523
CC1/2	0.999 (0.834)
CC*	1 (0.954)
R-work	0.16 (0.27)[d]
R-free	0.17 (0.29)[e]
Number of non-hydrogen atoms	3302
macromolecules	2928
ligands[f]	17
water	357
Protein residues	356
RMSD[g] **(bonds)**	0.006
RMSD[g] **(angles)**	1.07
Ramachandran favored (%)	97
Ramachandran outliers (%)	0
Average B-factor (Å)	31.10
Macromolecules (Å)	30.00
Ligandsf (Å)	44.60
Solvent (Å)	39.10

[a] Statistics for the highest-resolution shell are shown in parentheses.

[b] Mean [$I/\sigma(I)$] is the average of the relation between the intensity of the diffraction and the background.

[c] $R_{meas} = \{\Sigma_{hkl} [N/(N-1)]^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle|\} / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and N is the multiplicity of reflection *hkl*.

[d] R-work = $\Sigma_{hkl} \{|[F_{obs}(hkl)]| - |[F_{calc}(hkl)]|\} / \Sigma_{hkl} [F_{obs}(hkl)]$, where $F_{obs}(hkl)$ and $F_{calc}(hkl)$ are the structure factors observed and calculated, respectively.

[e] R-free corresponds to R_{factor} calculated using 2% of the total reflections selected randomly and excluded during refinement.

[f] Ligands: glycerol, ethylene glycol, ethanol.

[g] RMSD is the root mean square deviation.

doi:10.1371/journal.pone.0123463.t001

far-UV CD (circular dichroism) spectra were protein-concentration dependent; in that range, we did not observe variation in any spectral parameter. The cell path-length was 0.1 cm. Every experiment was repeated three times with new samples.

Fluorescence

Spectra were collected on a Cary Eclipse spectrofluorometer (Agilent) interfaced with a Peltier system at 25°C. A 1-cm-path-length quartz cell (Hellma) was used. The proper blank solutions were subtracted in all cases. Samples were prepared by taking the corresponding amount of a concentrated stock solution of $Erb1_{518\text{-}586}$ protein to yield a final protein concentration of 1μM. Spectra of $Erb1_{518\text{-}586}$ in aqueous solution (pH 7.0, 10mM Hepes and 0.150M NaCl) were acquired by excitation at 280 and 295nm; the emission spectra were collected between 300 and 400 nm. The excitation and emission slits were set to 5 nm, and the response was set to 1 nm.

NMR spectroscopy

The 1D-^{1}H-NMR and 2D- (^{1}H, ^{15}N) HSQC experiments of $Erb1_{518\text{-}586}$ were acquired at 25°C on a Bruker Avance DRX-500 spectrometer (Bruker GmbH, Germany), equipped with a triple resonance probe and z-pulse field gradients. Processing of spectra was carried out with TOPSPIN software. Spectra were calibrated with external TSP for the ^{1}H dimension, and for the ^{15}N dimension as described [25]. All spectra were acquired in phosphate buffer (pH 7, 50 mM).

1D-1H-NMR experiments. Protein concentration was 80 μM. Water was suppressed with the WATERGATE sequence [26]. A number of 512 scans were acquired with a spectral width of 12 ppm, with 16 K data points in the time domain. The data matrix was zero filled to 32 K during processing.

2D 15N-HSQC experiments. Spectra of ^{15}N-labeled $Erb1_{580\text{-}586}$ were acquired in the phase-sensitive mode; frequency discrimination in the indirect dimension was achieved by using the echo/antiecho-TPPI method. The 2D ^{15}N-HSQC experiment [27] was acquired with 4K data points in the ^{1}H dimension and 200 scans in the ^{15}N axis. The spectral widths were 15 and 35 ppm in the ^{1}H and ^{15}N dimensions, respectively. Carrier frequency of ^{1}H was set at 4.8 ppm and that of ^{15}N was 120 ppm. Water was suppressed with the WATERGATE sequence [26]. The concentration of $Erb1_{580\text{-}586}$ was 200 μM.

In vitro RNA binding assays

Poly(U)-agarose beads binding. Approximately 20μl of the polyuridylic acid-agarose (polyU) were used by assay. The beads were equilibrated 5 times with 500μl of reaction buffer each (50mM Tris pH 8; 100mM NaCl; 5mM $MgCl_2$; 3mM DTT; 0.1% Triton-X100 and 0.1 mg/ml BSA) then 500μl of the sample containing 200μg of protein diluted in the reaction buffer were loaded on the beads and incubated at room temperature for 30 minutes. Unbound protein was removed by washing the beads five times with wash buffer (50mM Tris pH 8; 100mM NaCl; 5mM $MgCl_2$; 3mM DTT; 0.1% Triton-X100). The bound fraction was eluted by addition of 30μl of 6x loading buffer and boiling for 5 minutes at 95°C. The samples were then analyzed on 10% polyacrylamide gel. The saturation was studied by incubating the protein sample with 0.1mg/ml or 1mg/ml of free polyuridylic acid for 30 minutes before it was loaded onto the beads. 1mg/ml of heparin was added to the protein sample prior to the poly(U) agarose binding in order to estimate the strength of the interaction.

Determination of the ChErb1432-801-polyU binding constant. Spectra of isolated $ChErb1_{432\text{-}801}$ and of the complexes of $ChErb1_{432\text{-}801}$/polyU, at different polyU concentrations, were acquired by excitation at 280 or 295 nm; the emission was collected between 300 and 400 nm. The excitation and emission slits were set to 5 nm, and the response was 1 nm.

Binding experiments of the polyU to ChErb1$_{432\text{-}801}$ were carried out as described [28]. Briefly, increasing amounts of polyU, in the range 0.5–5 μM, were added to a solution containing 1.5 μM of ChErb1, in 20 mM Hepes (pH 7.5), 150 mM NaCl, 5% glycerol and 2 mM ß-mercaptoethanol; the fluorescence was measured at 25°C, after preparing the samples. The apparent dissociation constants of the ChErb1-polyU complex, KD, was calculated by fitting the fluorescence change in intensity [29], versus the concentration of added polyU to:

$$F_{meas} = F + \Delta F_{\max}((([polyU] + [ChErb1] + K_D) - \sqrt{([polyU] + [ChErb1] + K_D)^2 - 4[ChErb1][polyU]}) \quad (1)$$

where [polyU] is the concentration of ribonucleic acid; [ChErb1] is the concentration of the protein; Fmeas is the measured fluorescence parameter at each concentration of added polyU; ΔFmax is the change in that parameter, when all ChErb1$_{432\text{-}801}$ is forming the complex; and F is the fluorescence parameter of the mixture of the complex Fitting to Eq (1) was carried out with Kaleidagraph (Abelbeck software).

Biolayer interferometry. Dissociation constant (KD) was determined more accurately by BioLayer Interferometry using BLItz system (ForteBio). A sample containing 50μg/ml of 15 nucleotide-long 5'-biotinylated polyU from Sigma-Aldrich was immobilized on Streptavidin biosensors (Forte Bio) previously hydrated with sample buffer (50mM Hepes pH 7.5; 150mM NaCl; 5% glycerol and 2mM β-mercaptoethanol). Increasing amounts of ChErb1$_{432\text{-}801}$ (0μM; 2μM; 5μM; 7.5 μM; 15 μM and 23μM) were used in association and dissociation steps. Curve fitting of triplicates and K_D calculation were carried out with BLItz Pro 1.2 software.

Results

Crystal unit cell suggests crystallization of a fragment produced by proteolytic cleavage

Initial analysis of the diffraction data showed that the asymmetric unit volume was 125433.9Å^3, not large enough to accommodate Erb1/Nop7 complex nor Erb1 alone. For Nop7 alone the Matthews coefficient and solvent content were 1.78Å^3/Da and 31% respectively, lower than expected [30]. In addition, the Xtriage module from Phenix estimated 470 residues as the most probable in the asymmetric unit which led us to investigate whether the full length Nop7 or a fragment of any of the two components had been crystallized.

In order to confirm if the crystals from initial screenings contained Nop7 or a fragment of Erb1, 15 crystals were analyzed by SDS-PAGE and, after staining with Coomassie Blue, only a single faint band of approximately 45 kDa could be observed on the gel. When the stability of Nop7, Erb1 and the Nop7/Erb1 complex was assayed, it was seen that after 24h of incubation at 4°C Erb1 started to show a clear pattern of degradation that was visible even upon binding to Nop7 and led to the apparition of the 45kDa band as seen for the crystals (Fig 1a). The mass-spec analysis confirmed that the lower MW band corresponded to the C-terminal region of Erb1 and contained the whole β-propeller domain of the protein.

The carboxy-terminal domain of Erb1 is a 7 bladed β-propeller

The solved structure of the C-terminal domain of Erb1 has confirmed that it folds into a seven-bladed β-propeller as previously predicted [1]. The blade organization and the nomenclature are shown on the Fig 1b and 1c. Structure analysis led us to slightly extend exact boundaries of the domain (residues 427–807) towards the N-term, when compared to the sequence-based prediction (residues 435–807) because the strand D (outermost) of blade 7 (most C-terminal)

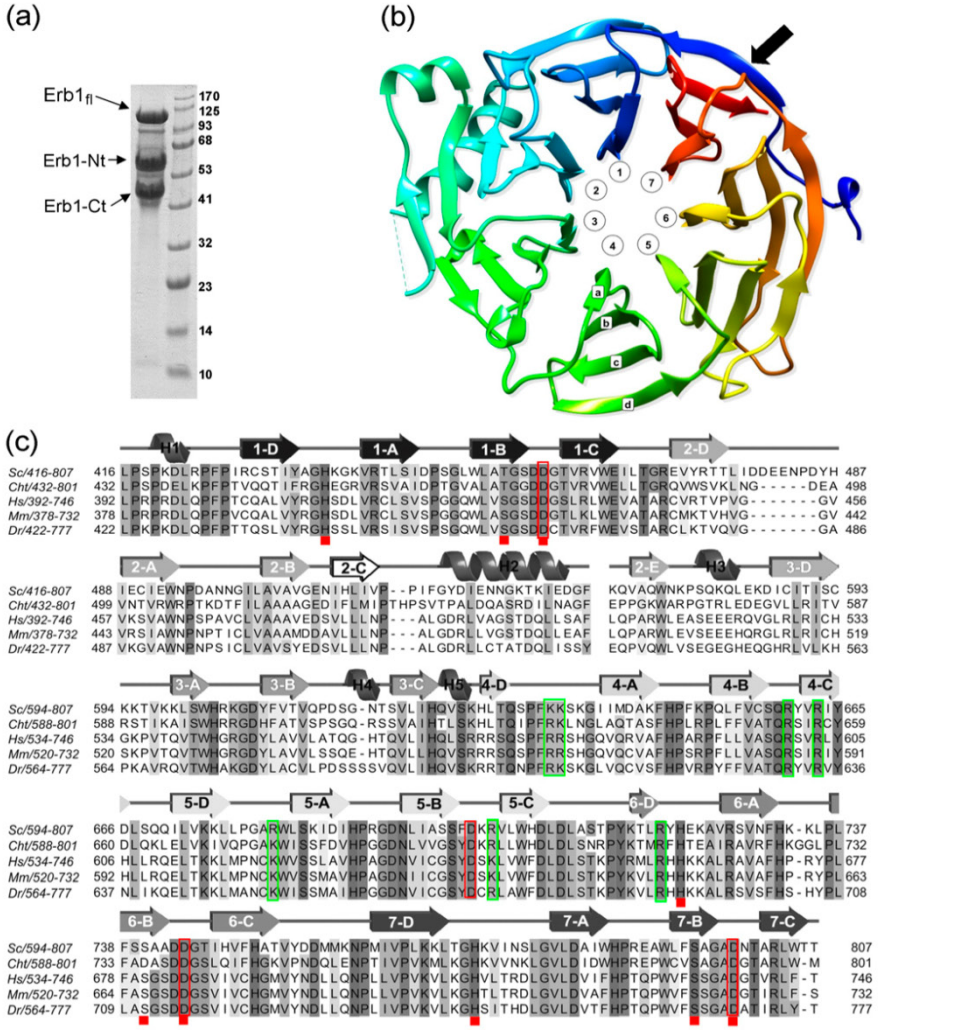


Fig 1. Erb1 degradation and β-propeller general features. (a) Coomassie—stained SDS-PAGE showing a severe degradation pattern of full-length Erb1 when incubated at 4°C overnight. The N-terminal and C-terminal degradation products are marked. MW in kDa is shown next to the ladder. (b) Ribbon representation of the β-propeller domain of Erb1 seen from the top (by convention the top face is described as the one that contains loops B-C and D-A). The blades numbering is counterclockwise and the β-strands nomenclature follows as shown for blade 4. Black arrow indicates the Velcro-like closure of the domain. (c) Sequence multi alignment of the carboxy-terminal domain of Erb1/Bop1 from *Saccharomyces cerevisiae* (Sc), *Chaetomium thermophilum* (Cht),

Homo sapiens (Hs), *Mus musculus* (Mm) and *Danio rerio (Dr)*. For clarity only the residues present in the final pdb model are shown. Conserved amino acids are marked with shadows. Secondary structure assignment is shown on the top of the alignment. Numbers of β-strands correspond to the WD repeats of the protein. Red rectangles mark conserved Asp in B-C loops and red squares indicate residues forming His-Thr/Ser-Asp triads. Green rectangles show basic residues that form a putative RNA-binding site.

doi:10.1371/journal.pone.0123463.g001

is actually formed by residues 427–433 which participate in the "velcro-like" closure of the domain (Fig 1b). Curiously, none of the WD repeats contains the eponymous WD motif but they rather present HD/YD dipeptides at the end of the strand C. Similarly to other WD40 domains, no clear pattern can be observed between the sequences of the repeats, although the hydrophobic core of the domain is well conserved and forms intra-molecular interactions necessary for proper folding. One of the most conserved features common for the β-propeller folds is the presence of a non-variable Asp in the loops that connect strands B and C of each blade. This residue is involved in stabilization of the domain as it forms a triad with a conserved His from the GH motif and a Ser/Thr residue placed in strand B. In Erb1, five of B-C loops contain an Asp residue but only four of them are truly conserved (red boxes Fig 1c) and the triad appears only in blades 1, 6 and 7 (Fig 2 and red squares in the alignment from Fig 1c). In loop 2B-2C there is a glutamic acid (Glu508) and in loop 4B-4C a glutamine (Gln659), both are conserved and establish a network of interactions that stabilizes the folding but is not similar to the canonical Asp-His-Ser/Thr triad. In the third blade, non-conserved Asp615 seems to be important for inter-blade interactions as it establishes hydrogen bonds with neighboring residues (Lys595 and Ile641), but it is rather to be substituted by a Gln in higher eukaryotes (Gln556 in human Bop1). It has been proposed that the conservation of the Asp residue could be correlated with the interacting His which means that if a WD repeat lacks the Asp it would also loose the His during evolution [31]. This seems to be true for Erb1 as the His from D-A loop is only

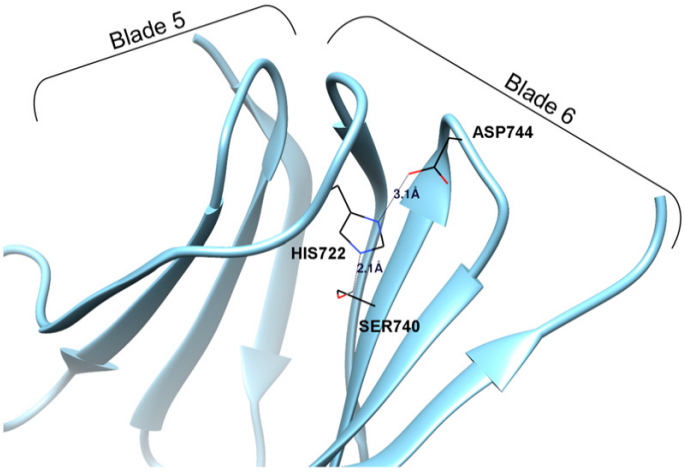


Fig 2. A canonical Asp/His/Ser triad necessary for correct blade organization. Conserved residues participating in the crucial hydrogen bond formation are shown and labelled. The distances between the atoms directly involved in H-H bonding are also represented.

doi:10.1371/journal.pone.0123463.g002

conserved in the three repeats that form the triad with Asp. At last, Asp701 from B-C loop of blade 5 is fully conserved but the His is not present, thus the triad cannot be formed. Instead, this residue altogether with conserved Arg703 makes electrostatic and hydrogen-bonds that stabilize the fifth blade.

The overall architecture of Erb1Ct superposes well with other 7-bladed propellers and is structurally highly similar to the β-propeller domain of the transducin-like enhancer protein1 (PDB: 2CE9) with the RMSD (Root-Mean-Square Deviation) score of 2.2Å as calculated by Dali Server [32]. The main difference between these characteristic propeller folds lies in the presence of loops that connect β-strands and contribute to the functional specialization of the domain by creating surfaces that will enable specific protein-protein interactions. Regarding Erb1, some of these loops are quite long and might be very flexible as reflected by high temperature factors corresponding to the residues 478–488 that connect strands 2D and 2A, nevertheless when compared to Erb1/Bop1 in higher eukaryotes, a clear tendency to loop shortening may be observed (Fig 1c). Interestingly, we detected a big variation regarding the conservation of the loops that connect the blades. In general, the loops between strands A-B (bottom side of the propeller) are more variable in sequence, whereas those on the top face, connecting strands B and C (which contain Asp residues highlighted in Fig 1c), present a higher degree of conservation thus indicating an important region for the β-propeller stability or a conserved protein-protein interaction surface.

From a crystallographic point of view, unlike in several crystal structures of β-propellers [33–35], the extensive top or bottom areas of the domain do not form many contacts with symmetry-related molecules but the interactions are rather maintained laterally through the outermost loops and strands (Fig 3a). Manual inspection of the unit cell and analysis by Pisa Server [36] showed that a single monomer contacted six other molecules making three small interfaces involved in crystal packing. Two of them arose in the bottom part of the propeller where the alpha-helix H2 orientates in proximity of blades 1 and 7 from one symmetry related monomer and the loops 5B-5C and 6D-6A from another propeller (Fig 3b). The third area corresponds to the long 6C-7D loop which forms an important extension and introduces Asp757, Met758 and Met759 into a cavity formed between Phe635, Tyr665 and Gln670 from blades 3 and 4 (Fig 3c). Both the loop and the cavity are well conserved within Erb1 family and may constitute additional elements involved in protein binding. The central axis of the propeller is filled with solvent (water) molecules.

The β-propeller domain of Erb1 contains a long insertion within blade 2

Undoubtedly, the most distinctive feature of the core of the β-propeller in this study is the blade 2 which due to an insertion contains five, and not four, β-strands and shows a protrusion attributed to two α-helices (H2 and H3 on Fig 1c). Electron density map allowed us to trace and build model for residues 515–534 and 571–594, being the rest of the insertion unmodeled. This missing part seems to be Fungi specific as it becomes much shorter in higher eukaryotes (Fig 1c). Helix H2 (residues Tyr520-Asp532) appears between strands 2C and 2E and is attached to the base of the propeller (Fig 4a). In general, the sequence of the helix is poorly conserved, but it contains two invariable non-polar residues: Ile522 that makes hydrophobic contacts with the backbone of 2A-2B loop and Ile530 which interacts with the C-terminal fragment of strand 3D. The β-sheet corresponding to this blade is formed by strands 2A, 2B, 2C, 3D and 2E clearly indicating an alteration of a standard WD40 pattern (Fig 4b). Whereas strand 3D unambiguously indicates the beginning of WD3, the sequence of WD repeat 2 does not show any significant conservation but still contains strategic residues that allow formation of hydrophobic and electrostatic interactions with neighboring blades. Initial sequence-based analysis suggested that

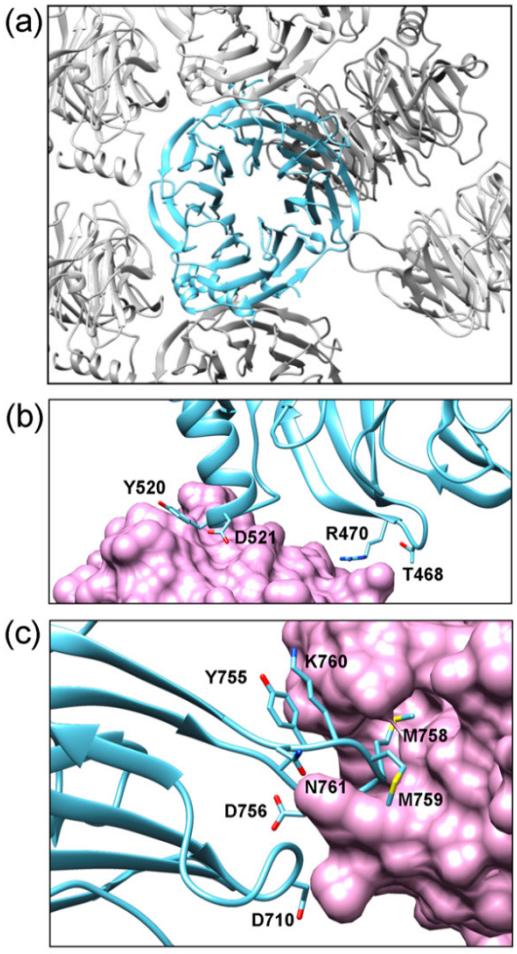


Fig 3. Analysis of crystal packing. (a) Overview of crystal contacts of Erb1 monomer (blue) with symmetry related molecules (grey) shows that the top and bottom areas of the propeller are not involved in crystallographic interactions. (b) Helix H2 interacts with symmetrically related molecule (shown in pink). (c) 6C-7D loop penetrates deeply into a conserved cavity of another monomer (in pink). The residues directly involved in crystal packing are labelled in (b) and (c).

doi:10.1371/journal.pone.0123463.g003

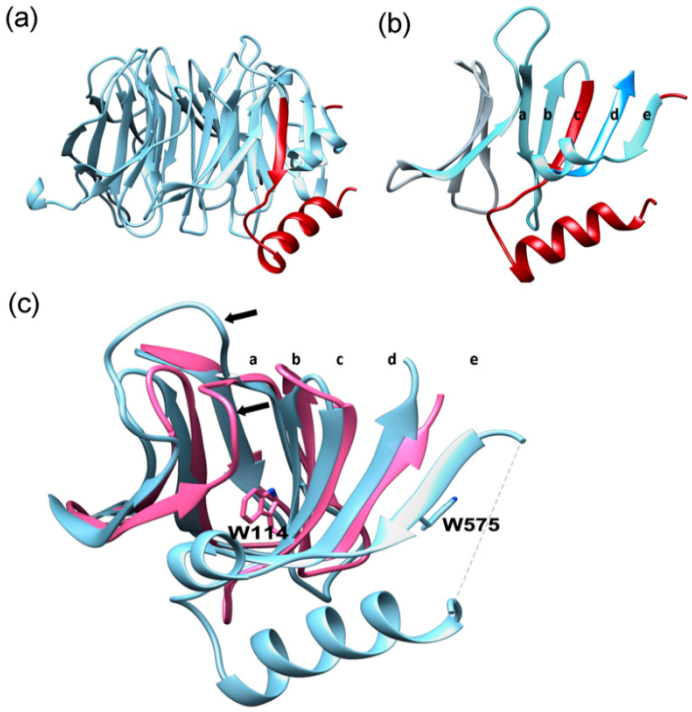


Fig 4. Insertion within WD repeat 2. (a) The insertion (red) forms an important protrusion on the bottom of the domain. (b) Position of the insertion (red) in the context of the second blade only. Residues corresponding to WD repeat 2 are represented in light blue and the strand D of WD repeat 3 is shown in dark blue. (c) Comparison of blades 1 and 2 of WDR5 from *H. sapiens* (PDB:4CY2; pink) and Erb1 (blue). Side chains of conserved tryptophan corresponding to the strand C (in canonical WD repeats) are shown for both proteins. Black arrows indicate the position of 2D-2A loops. The letters in (b) and (c) indicate the position of each strand.

doi:10.1371/journal.pone.0123463.g004

between WD repeats 2 and 3 there was an approximately 80-residue long segment which did not contain any WD pattern. Surprisingly, when we aligned the sequence of Ct domain of Erb1 with other non-Erb1/Bop1 β-propeller-containing proteins we could clearly see that Trp575 from strand 2E corresponded to Trp residue from WD dipeptide that typically appears in strand C (as in human WDR5 protein, PDB: 2H14) (Fig 4c) [37]. This fully conserved residue establishes important hydrophobic interactions with Ile592 from strand 2D and His629 located in 3D that are likely to be required for a proper conformation and attachment of the insertion to the side of the blade 2. We conclude that from an evolutionary point of view strand E corresponds to strand C from a canonical blade although displaced, in the second blade of Erb1, by a segment containing 2C-loop-H2. This insertion produced an important reorganization of the whole blade, altering the position and function of Trp-Asp dipeptide (Trp-Asn in this case). As a result, the second blade lacks the important Trp residue at the end of strand 2C that would

guarantee correct approach between blades 1 and 2. We observe that in this case there is a different interaction network, conserved in Erb1/Bop1 family but not in other WD repeat-containing proteins, that involves strand 2D from blade 1 and a short α-helix, H3, from blade 2 (Gln580-Lys585). This helix inserts between strands 2E and 3D and possesses two non-conserved lysine residues (Lys581; Lys585) that interact with loop 2D-2A through hydrogen bonds. In consequence of this arrangement, α-helix H3 forms a lid that orientates close to a very hydrophobic area in blade 2 created by a segment of well conserved polar residues from strand 2B. It is important to keep in mind that loop 2D-2A is quite flexible and its vertical orientation makes the whole interface between blades 1 and 2 more opened when compared with the gaps between other blades which are completely covered by D-A loops (Fig 4c).

The isolated insertion, Tyr518-Asp586 (Erb1$_{518\text{-}586}$), is disordered in solution

We studied the conformational propensities of the isolated insertion, the fragment comprising residues Tyr518-Asp586 of Erb1, in solution. Fluorescence spectrum showed the maximum at 345 nm, suggesting that the sole Trp, Trp575, was highly solvent-exposed (Fig 5a). Furthermore, its far-UV CD spectrum had an intense minimum at 200 nm, and a shoulder at 222 nm (Fig 5b), which suggests the presence of helical or turn-like conformations. Those structured conformations, however, were not stable enough, since thermal denaturations followed by far-UV CD did not show any sigmoidal behavior (data not shown). The 1D-^{1}H–NMR spectrum had all the amide protons clustered between 8.0 and 8.5 ppm. Moreover, the indole proton appeared at 10.19 ppm, close to the expected value of a random-coil indole moiety [38], and thus further supporting the fluorescence results. In the 1D-^{1}H-NMR spectrum, there was no downfield shifted H_{α} protons (suggesting the absence of β-strands); however, there were a few upfield shifted methyl protons, which are probably close, in the primary structure, to some of the aromatic groups (Fig 5c). Finally, the 2D-^{15}N-HSQC experiments (Fig 5d) had a very narrow dispersion in the amide proton region of 0.4 ppm that is centered around the random-coil chemical shift values for NHs (between 8.2 and 8.4 ppm), confirming the 1D-NMR results. In addition, there was a lower number of cross-peaks than we should expect (79); the absence of the peaks is either probably due to chemical exchange broadening or rapid amide exchange, as expected in natively unfolded proteins [39].

Evolutionary conserved clusters on the surface of the propeller form putative ligand binding pockets

The best known and most obvious function of β-propellers is their capacity to establish multiple interactions with other proteins. The intrinsic rigidity and the shape of the domain create three well defined zones where the binding partner can attach: the top, the bottom and the circumference of the propeller. We searched for conserved residues on the surface of Erb1Ct which could indicate a region important to preserve protein interaction interface. There is a very clear division between a poorly conserved area, that includes blades 1, 2 and the upper part of blade 3, and a less variable surface of blades 4, 5, 6 and 7. In the bottom part, between blades 3 and 4 we identified a well conserved pocket which is a good candidate as a possible place of association with a binding partner.

In addition, it has been proposed that the central channel of WD domains could work as a scaffold that adapts for recognition of different ligands through side-chains of three residues from each blade: the one right before the strand A (A-1), the one just after strand B (B+1) and the second residue in the strand A (A2), thus making this portion of the propeller an universal but variable binding motif. When we inspected these positions in Erb1Ct a strong conservation,

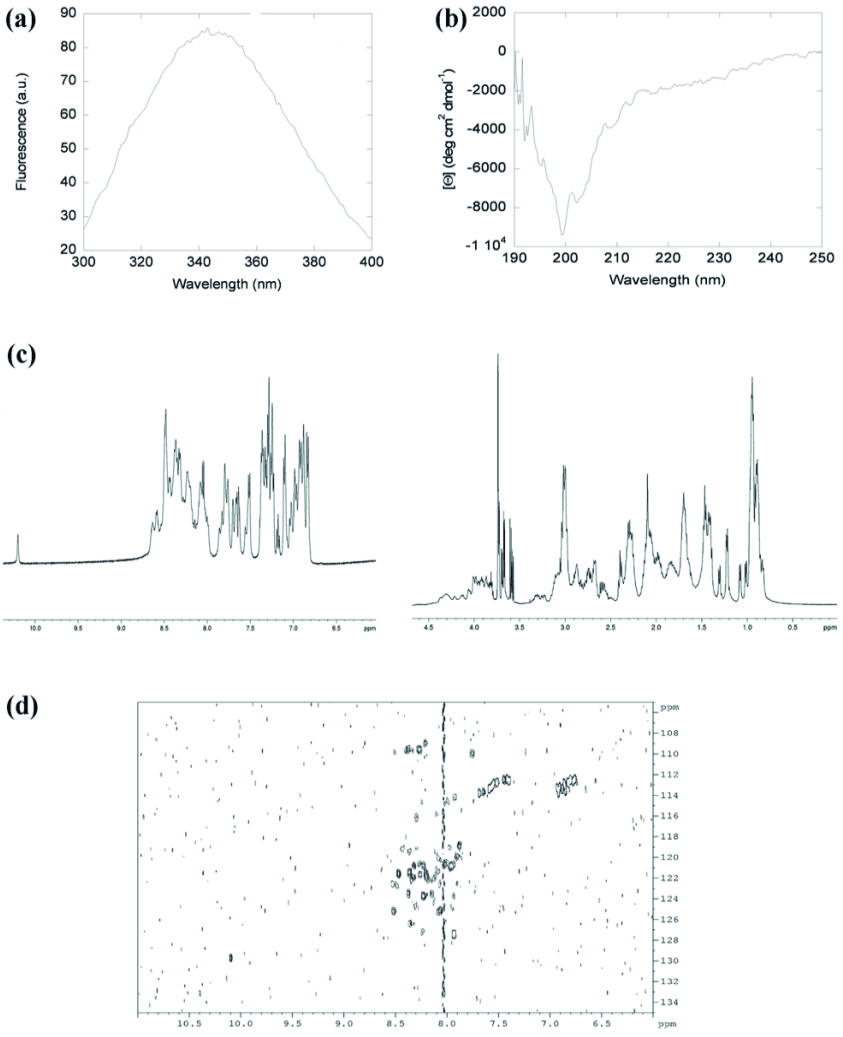


Fig 5. Erb1 insertion is disordered in solution. (a) Fluorescence spectrum of $Erb1_{518-586}$ obtained by excitation at 280 nm. (b) Far-UV CD spectrum of $Erb1_{518-586}$. (c) Amide (left) and alkyl (right) regions of the 1D-^{1}H- NMR spectrum of $Erb1_{518-586}$. (d) ^{15}N-^{1}H HSQC spectrum of $Erb1_{518-586}$.

doi:10.1371/journal.pone.0123463.g005

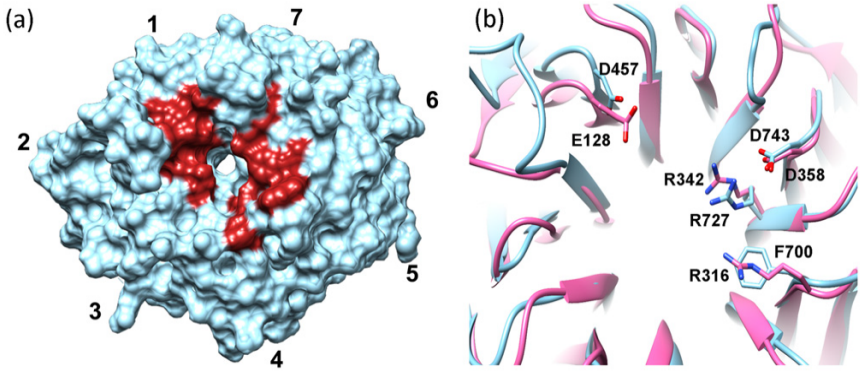


Fig 6. Top face of the β-propeller contains "hot spot" residues. (a) Top face of the propeller showing the position (in red) of the residues most likely to participate in macromolecular interactions as predicted by WDSP server (http://wu.scbb.pkusz.edu.cn/wdsp/). (b) Superposition of Lis1 (pink, PDB: 1VYH) with Erb1 β-propeller. The side chains of the conserved amino acids are shown and labeled.

doi:10.1371/journal.pone.0123463.g006

especially in blades 1, 5, 6 and 7, was observed. Our findings were confirmed by WDSP web server which predicted hot-spot residues on the surface that were likely to be responsible for high-affinity interactions with other proteins [40] (Fig 6a). Moreover, those conserved positions seem to be related to Erb1/Bop1 function because they vary when compared to seven-bladed propellers from other families. Nevertheless, three of these superficial conserved amino acids, Asp457, Arg727 and Asp743 are also invariable in another family of WD repeat-containing proteins called Lis1 where they were shown to be involved in recognition of other macromolecules [41] (Fig 6b).

The carboxy-terminal domain of Erb1 binds RNA in vitro

Erb1 is known to bind rRNA as shown by UV-crosslinking experiments. It is possible that the β-propeller is involved in such a binding due to its highly positive charge (Fig 7). We used the PatchFinder Plus algorithm [42] to identify the biggest positive patch on the surface of the domain. Indeed, as seen for the electrostatic surface analysis, the tool found a big region of basic residues on the surface that included five blades and the entrance to the central channel on the top face of the propeller. Curiously, one of the few known structures of a WD domain bound to a nucleic acid is the one of DDB1-DDB2 complex with a DNA chain binding the protein through a cavity formed by the arginine and lysine residues oriented around the central tunnel of the propeller [43]. In Erb1 this area also contains well conserved basic amino acids: Arg441, Lys598 and Arg727.

Initially, we checked the affinity of the β-propeller of Erb1 for RNA in vitro using poly(U) agarose beads. Our first attempts to perform in vitro binding assays failed because the C-terminal domain of Erb1 (residues 416–807) from yeast expressed poorly in *E. coli* and rapidly degraded during purification. We decided to try whether the same domain from a thermophile would be stable enough to carry out the experiments. Finally, a truncated Erb1 from *Chaetomium thermophilum*, comprising residues 432–801 ($ChErb1_{432-801}$), was used for the assays due to its enhanced stability and higher expression levels in *E. coli*.Since the sequence of the domain is well conserved between *S. cerevisiae* and *C. thermophilum*, including the basic residues

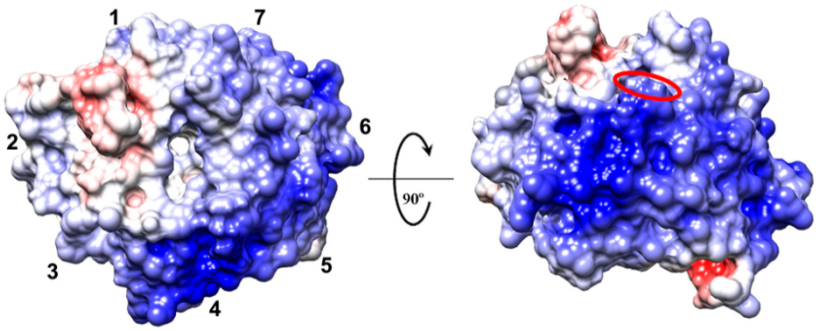


Fig 7. Surface of Erb1 β-propeller is positively charged. Surface representation of the electrostatic potential of the domain (from red (-10) to blue (+10)k_b T e_c^{-1}). The top face is shown on the left and the most positively charged area formed by blades 4 and 5 is visible on the right panel. The red oval indicates the position of Trp682.

doi:10.1371/journal.pone.0123463.g007

from the putative RNA-binding area (shown with green boxes in Fig 1c), we consider $ChErb1_{432-801}$ to be suitable for validation of our findings based on $Erb1_{416-807}$ structure from yeast. As shown in the Fig 8a the propeller appeared in the eluate from poly(U) beads. To investigate whether the interaction occurred through a well-defined surface that could be saturated, we incubated the protein with free poly(U) before it was loaded on the beads. The amount of the protein that could stably bind the poly(U) beads decreased in presence of 0.1mg/ml and 1mg/ml free poly(U), indicating saturation of the binding site (Fig 8b). The binding affinity was good because it was still detectable even when heparin was added to the reaction mixture (Fig 8c). Fluorescence experiments were carried out in order to estimate the binding affinity. Emission spectrum of intact $ChErb1_{432-801}$ shows a maximum wavelength at 340 nm (Fig 8d), indicating that the tryptophans in the structure are partially buried, as shown in the X-ray structure. Therefore, if the RNA binding occurred in the proximity of the indole moiety, we should expect changes in the fluorescence spectra upon addition of polyU, and therefore we could use these changes to measure the affinity constants by using Eq (1). We acquired spectra at growing concentrations of polyU (Fig 8e), and the changes in the fluorescence emission spectra at 315 nm allowed us to determine the constant (similar changes were observed at other wavelengths either by excitation at 280 or 295 nm). The apparent K_D of the complex was 3 ± 2 μM. The fact that an exposed tryptophan (Trp682, red oval in the Fig 7a) is at the vicinity of the positively charged stretch (likely to participate in RNA binding), provides a good indication that the interaction takes place through the proposed area and explains the change in fluorescence upon binding of the nucleic acid. We also calculated binding affinity between $ChErb1_{432-801}$ and a 15 nucleotide-long biotinylated poly(U) using biolayer interferometry. The dissociation constant (K_D) was 0.17 μM (Fig 8f).

Discussion

Our initial studies on Erb1 show that it is an unstable protein that rapidly undergoes proteolytic degradation when overexpressed in *E. coli* even upon binding to its functional partner Nop7. Since the interaction of Nop7 and Erb1 is strong (they co-elute in size exclusion chromatography), the proteolytic cleavage occurs somewhere between Nop7 binding site and the ß-

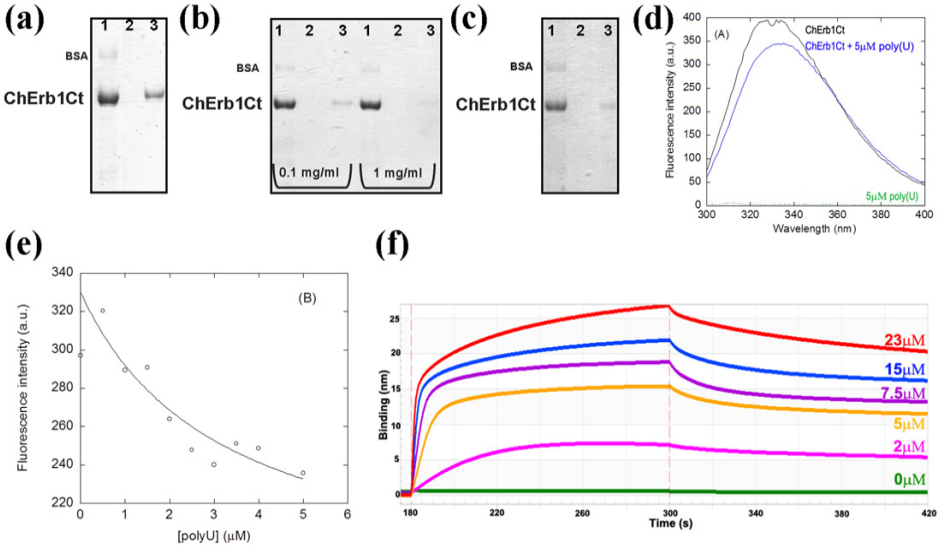


Fig 8. ChErb1Ct (residues 432–801) binds RNA *in vitro*. (a) Coomassie-stained SDS-PAGE showing the binding of Erb1 β-propeller from *Ch. thermophilum* to polyU agarose beads. (b) The saturation of the binding is visible upon addition of 0.1 mg/ml or 1 mg/ml of free polyuridilic acid. (c) The binding is still detectable upon addition of 1 mg/ml of heparin to the binding buffer. (a) (b) and (c) 1: Input, 2: Wash, 3: Elution; (d) Fluorescence spectra of $ChErb1_{432-801}$ alone (in black) and with 5μM poly(U) (blue) obtained by excitation at 280 nm. The spectra were acquired at 25°C with 1.5 μM of $ChErb1_{432-801}$. The green line at the bottom of the spectra is the emission spectra of polyU at a concentration of 5 μM. (e) Titration curve obtained from the emission fluorescence intensity at 315 nm. The line is the fitting to Eq (1). (f) Association and dissociation curves of 15nt-poly(U) with different concentrations of $ChErb1_{432-801}$ measured by biolayer interferometry.

doi:10.1371/journal.pone.0123463.g008

propeller domain. In fact, Bop1, mammalian ortholog of Erb1, was predicted to possess a PEST (proline, glutamic acid, serine and threonine rich sequence) element in its central region that overlaps with the predicted Nop7 binding sequence [2]. The SitePrediction tool [44] confirmed the presence of a conserved PEST sequence in Erb1 in the same position as for Bop1. PEST motifs are clusters of charged amino acids that form flexible and solvent exposed loops that are linked to protein destabilization and degradation in eukaryotic systems [45]. Erb1 protein used in our study was expressed in bacteria where no PEST-related protein degradation mechanism has been described but it is possible that, despite Nop7 binding, the region is easily accessible for *E. coli* proteases, leading to a rapid Erb1 cleavage. It is known that Nop7/Erb1 complex is only stable when bound to Ytm1 and the predicted Ytm1 binding site is located in proximity of the PEST sequence and the β-propeller domain [6]. The interdependent stability of these factors could mark one of the many control checkpoints during ribosome assembly, in which the proteasome pathway might be involved.

The structure of the carboxy-terminal domain of Erb1 from *Saccharomyces cerevisiae* reveals that it adopts the classical seven-bladed β-propeller conformation but, at the same time, it presents a series of features that makes it distinguishable from the rest of the WD repeat-containing proteins. The sequence of the propeller is well conserved within the members of Erb1/

Bop1 family but is quite dissimilar to other proteins with the same domain resulting in another good example of how the propeller fold is maintained throughout the evolution in spite of poor sequence conservation.

Even though, there is no clear evidence for Erb1Ct function in the cell, it is likely to be involved in mediating protein-protein or protein-rRNA interactions in eukaryotic ribosome assembly. Search of possible binding interfaces showed that the β-propeller of Erb1 contains several regions that might bind various ligands. These patches are conserved and appear in areas that, in other WD40-containing proteins, were shown to participate in binding to different targets. Since the charge on the surface of the propeller suggests that it could bind RNA, we confirm this hypothesis in vitro proving that a good-affinity binding takes place between the C-terminal domain and polyuridylic acid. Although we could also observe interaction between the β-propeller and DNA (data not shown), from a functional point of view RNA binding seems more likely to occur in vivo. The specificity of the interaction might be provided by Nop7/Erb1 attachment site on the nascent ribosome which results in proximity of Erb1 C-term to rRNA. This binding could contribute to the stability of Nop7 sub-complex association with pre-ribosome. We identified a region close to the central channel opening, on the narrow side of the propeller, that includes few conserved residues that mediate protein-protein interactions and appear in the same position as in Lis1 protein, a platelet-activating factor that binds to dynein and other ligands through this conserved surface [41]. In addition, in several β-propellers the same region was shown to be responsible for binding to other proteins or ligands and it was proposed as one of the most common interaction sites within the domain [46].

Recently, it has been postulated that the WD40 proteins can also establish interactions by additional extensions or insertions that can occur between blades of the propeller [47]. It could be valid for Erb1 because, as we show here, it contains a long region that is introduced into WD2 and forms two visible α-helices and an additional strand. Residues in the propeller, outside the insertion, can contribute to the folding of the observed secondary motifs in this area. While the extra strand in blade 2 is probably necessary for structural stability of the inserted region and the propeller, the helices form an important protrusion that could serve as recognition motifs that selectively bind a ligand. In fact, the part that is not visible in the model (residues 535–570) —probably due to its flexibility and lack of secondary structure (as shown by the spectroscopic findings observed with the isolated region)—could become ordered and more rigid upon binding to a protein or nucleic acid. Since the top entrance of the tunnel and the helices from the insertions are located on opposite sides of the propeller, it is possible that the domain recruits two different binding partners at the same time (for a summary of the putative binding sites see Fig 9). It is also worth to note that C-D loop in blade 6 noticeably projects out of the bottom plane of the propeller in order to penetrate into a cavity of a symmetry-related molecule providing additional segments capable of interacting with other proteins. Moreover, if we take into consideration that Erb1 carries Nop7 and Ytm1 binding motifs towards its N-terminus it is plausible that the full length protein can establish multiple interactions within pre-ribosomal particles acting as a scaffold that plays an essential role in 60S maturation.

Taken together, our observations demonstrate that the C-terminal domain of Erb1 is likely to form part of a protein complex and is capable of binding RNA unspecifically through a defined surface. Although Granneman *et al.* showed that Erb1 bound to a specific region of 25S rRNA [9], they did not define the exact portion of the protein involved in the interaction. Our findings clearly suggest that the β-propeller domain might mediate this union. The size of the domain, its degree of conservation and the ability to bind nucleic acids supports the idea that Erb1 requires its carboxy-terminus to associate with other factors. Further functional studies need to be carried out in order to review the role of Erb1Ct in ribosome biogenesis especially if its propensity to bind RNA is considered.

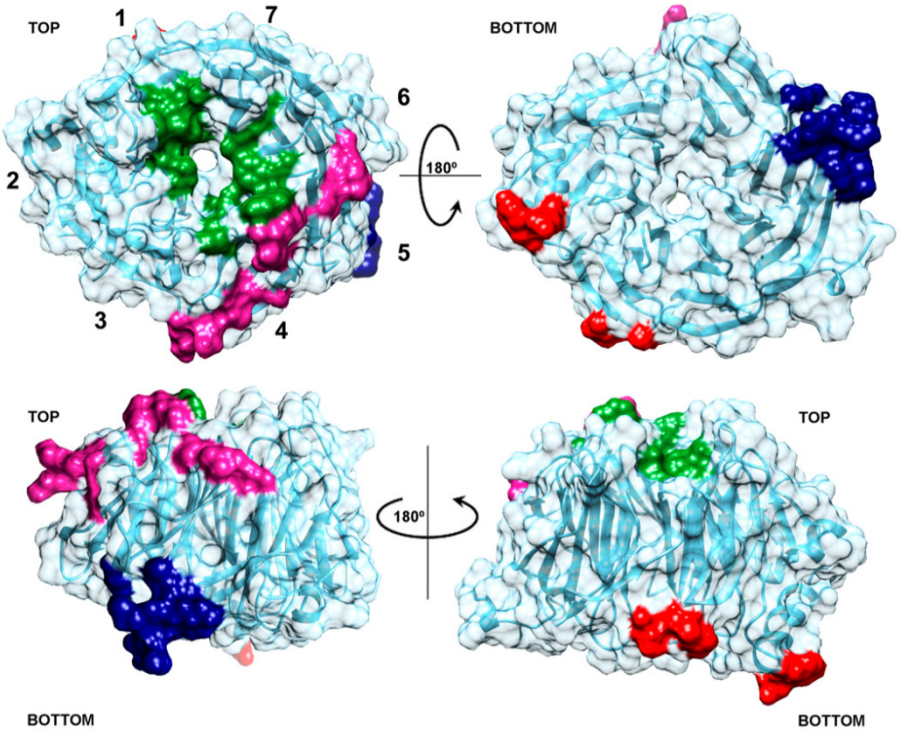


Fig 9. Summary of the areas of Erb1Ct that are more likely to mediate protein-protein or protein-RNA interactions. The region corresponding to the conserved hot-spot residues is shown in green, the positively charged stretch of amino acids that might interact with RNA is highlighted in magenta, navy blue and red areas correspond to the conserved parts of the domain that establish important crystal contacts with symmetrically related molecules.

doi:10.1371/journal.pone.0123463.g009

Acknowledgments

The authors would like to thank Diamond Light Source for beam-time (proposal mx8035), and the staff of beam-lines I03 and I24 for assistance with crystal testing and data collection.

Author Contributions

Conceived and designed the experiments: MW JB. Performed the experiments: MW JLN. Analyzed the data: MW JLN JB. Contributed reagents/materials/analysis tools: JLN JB. Wrote the paper: MW JLN JB.

References

1. Pestov DG, Stockelman MG, Strezoska Z, Lau LF. ERB1, the yeast homolog of mammalian Bop1, is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs. Nucleic Acids Res.

2001; 29: 3621–30. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=55883&tool=pmcentrez&rendertype=abstract PMID: 11522832

2. Strezoska Z, Pestov DG, Lau LF. Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. J Biol Chem. 2002; 277: 29617–25. doi: 10.1074/jbc.M204381200 PMID: 12048210
3. Strezoska Z, Pestov DG, Lau LF. Bop1 is a mouse WD40 repeat nucleolar protein involved in 28S and 5. 8S RRNA processing and 60S ribosome biogenesis. Mol Cell Biol. 2000; 20: 5516–28. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=86002&tool=pmcentrez&rendertype=abstract PMID: 10891491
4. Rohrmoser M, Ho M, Grimm T, Malamoussi A, Harasim T, Orban M, et al. Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. Mol Cell Biol. 2007; 27: 3682–94. doi: 10.1128/MCB.00172-07 PMID: 17353269
5. Killian A, Le Meur N, Sesboüé R, Bourguignon J, Bougeard G, Gautherot J, et al. Inactivation of the RRB1-Pescadillo pathway involved in ribosome biogenesis induces chromosomal instability. Oncogene. 2004; 23: 8597–602. doi: 10.1038/sj.onc.1207845 PMID: 15467761
6. Tang L, Sahasranaman A, Jakovljevic J, Schleifman E, Woolford JL. Interactions among Ytm1, Erb1, and Nop7 required for assembly of the Nop7-subcomplex in yeast preribosomes. Mol Biol Cell. 2008; 19: 2844–56. doi: 10.1091/mbc.E07-12-1281 PMID: 18448671
7. Sahasranaman A, Dembowski J, Strahler J, Andrews P, Maddock J, Woolford JL. Assembly of Saccharomyces cerevisiae 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing. EMBO J. Nature Publishing Group; 2011; 30: 4020–32. doi: 10.1038/emboj.2011.338
8. Sloan KE, Mattijssen S, Lebaron S, Tollervey D, Pruijn GJM, Watkins NJ. Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. J Cell Biol. 2013; 200: 577–88. doi: 10.1083/jcb.201207131 PMID: 23439679
9. Granneman S, Petfalski E, Tollervey D. A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. EMBO J. Nature Publishing Group; 2011; 30: 4006–4019. doi: 10.1038/emboj.2011.256 PMID: 21811236
10. Hölzel M, Rohrmoser M, Schlee M, Grimm T, Harasim T, Malamoussi A, et al. Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. J Cell Biol. 2005; 170: 367–78. doi: 10.1083/jcb.200501141 PMID: 16043514
11. Miles TD, Jakovljevic J, Horsey EW, Harnpicharnchai P, Tang L, Woolford JL. Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes. Mol Cell Biol. 2005; 25: 10419–32. doi: 10.1128/MCB.25.23.10419-10432.2005 PMID: 16287855
12. Bassler J, Kallas M, Pertschy B, Ulbrich C, Thoms M, Hurt E. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. Mol Cell. Elsevier Ltd; 2010; 38: 712–21. doi: 10.1016/j.molcel.2010.05.024 PMID: 20542003
13. Henras a K, Soudet J, Gérus M, Lebaron S, Caizergues-Ferrer M, Mougin a, et al. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci. 2008; 65: 2334–59. doi: 10.1007/s00018-008-8027-0 PMID: 18408888
14. Nal B, Mohr E, Silva M-I Da, Tagett R, Navarro C, Carroll P, et al. Wdr12, a mouse gene encoding a novel WD-Repeat Protein with a notchless-like amino-terminal domain. Genomics. 2002; 79: 77–86. doi: 10.1006/geno.2001.6682 PMID: 11827460
15. Talkish J, Zhang J, Jakovljevic J, Horsey EW, Woolford JL. Hierarchical recruitment into nascent ribosomes of assembly factors required for 27SB pre-rRNA processing in Saccharomyces cerevisiae. Nucleic Acids Res. 2012; 40: 8646–61. doi: 10.1093/nar/gks609 PMID: 22735702
16. Bradatsch B, Leidig C, Granneman S, Gnädig M, Tollervey D, Böttcher B, et al. Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. Nat Struct Mol Biol. Nature Publishing Group; 2012; 19: 1234–41. doi: 10.1038/nsmb.2438 PMID: 23142978
17. Leidig C, Thoms M, Holdermann I, Bradatsch B, Berninghausen O, Bange G, et al. 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. Nat Commun. Nature Publishing Group; 2014; 5: 3491. doi: 10.1038/ncomms4491 PMID: 24662372
18. Mushtaq A, Jamill A. Cloning of a β-glucosidase gene from thermophilic fungus Cheatomium thermophilum. Pakistan J Life Soc Sci. 2012; 10: 98–101.
19. Studier FW. Protein production by auto-induction in high-density shaking cultures. Protein Expr Purif. 2005; 41: 207–234. doi: 10.1016/j.pep.2005.01.016 PMID: 15915565
20. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66: 125–32. doi: 10.1107/S0907444909047337 PMID: 20124692

21. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. International Union of Crystallography; 2011; 67: 235–42. doi: 10.1107/S0907444910045749

22. Adams PD, Afonine P V, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66: 213–21. doi: 10.1107/S0907444909052925 PMID: 20124702

23. Long F, Vagin A a, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. Acta Crystallogr D Biol Crystallogr. International Union of Crystallography; 2008; 64: 125–32. doi: 10.1107/S0907444907050172 PMID: 18094476

24. Woody RW. Circular dichroism. Methods Enzymol. 1995; 246: 34–71. Available: http://www.ncbi.nlm.nih.gov/pubmed/7538625 PMID: 7538625

25. Alcaraz LA, del Alamo M, Barrera FN, Mateu MG, Neira JL. Flexibility in HIV-1 assembly subunits: solution structure of the monomeric C-terminal domain of the capsid protein. Biophys J. 2007; 93: 1264–76. doi: 10.1529/biophysj.106.101089 PMID: 17526561

26. Piotto M, Saudek V, Sklenár V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR. 1992; 2: 661–5. Available: http://www.ncbi.nlm.nih.gov/pubmed/1490109 PMID: 1490109

27. Bodenhausen G, Ruben DJ. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. Chem Phys Lett. 1980; 69: 185–189. doi: 10.1016/0009-2614(80)80041-8

28. Muro-Pastor MI, Barrera FN, Reyes JC, Florencio FJ, Neira JL. The inactivating factor of glutamine synthetase, IF7, is a "natively unfolded" protein. Protein Sci. 2003; 12: 1443–54. doi: 10.1110/ps.0303203 PMID: 12824490

29. Czypionka A, de los Paños OR, Mateu MG, Barrera FN, Hurtado-Gómez E, Gómez J, et al. The isolated C-terminal domain of Ring1B is a dimer made of stable, well-structured monomers. Biochemistry. 2007; 46: 12764–76. doi: 10.1021/bi701343q PMID: 17935356

30. Kantardjieff KA, Rupp B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci. 2003; 12: 1865–71. doi: 10.1110/ps.0350503 PMID: 12930986

31. Garcia-Higuera I, Gaitatzes C, Smith TF, Neer EJ. Folding a WD repeat propeller. Role of highly conserved aspartic acid residues in the G protein beta subunit and Sec13. J Biol Chem. 1998; 273: 9041–9. Available: http://www.ncbi.nlm.nih.gov/pubmed/9535892 PMID: 9535892

32. Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010; 38: W545–9. doi: 10.1093/nar/gkq366 PMID: 20457744

33. Cheng Z, Liu Y, Wang C, Parker R, Song H. Crystal structure of Ski8p, a WD-repeat protein with dual roles in mRNA metabolism and meiotic recombination. Protein Sci. 2004; 13: 2673–84. doi: 10.1110/ps.04856504 PMID: 15340168

34. Pickles LM, Roe SM, Hemingway EJ, Stifani S, Pearl LH. Crystal Structure of the C-Terminal WD40 Repeat Domain of the Human Groucho/TLE1 Transcriptional Corepressor. Structure. 2002; 10: 751–761. doi: 10.1016/S0969-2126(02)00768-2 PMID: 12057191

35. Wilson DK, Cerna D, Chew E. The 1.1-angstrom structure of the spindle checkpoint protein Bub3p reveals functional regions. J Biol Chem. 2005; 280: 13944–51. doi: 10.1074/jbc.M412919200 PMID: 15644329

36. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007; 372: 774–97. doi: 10.1016/j.jmb.2007.05.022 PMID: 17681537

37. Couture J-F, Collazo E, Trievel RC. Molecular recognition of histone H3 by the WD40 protein WDR5. Nat Struct Mol Biol. 2006; 13: 698–703. doi: 10.1038/nsmb1116 PMID: 16829960

38. Wüthrich K. NMR of Proteins and Nucleic Acids. 1986.

39. Dyson HJ. Expanding the proteome: disordered and alternatively folded proteins. Q Rev Biophys. 2011; 44: 467–518. doi: 10.1017/S0033583511000060 PMID: 21729349

40. Wu X-H, Wang Y, Zhuo Z, Jiang F, Wu Y-D. Identifying the hotspots on the top faces of WD40-repeat proteins from their primary sequences by β-bulges and DHSW tetrads. PLoS One. 2012; 7: e43005. doi: 10.1371/journal.pone.0043005 PMID: 22916195

41. Tarricone C, Perrina F, Monzani S, Massimiliano L, Kim M-H, Derewenda ZS, et al. Coupling PAF signaling to dynein regulation: structure of LIS1 in complex with PAF-acetylhydrolase. Neuron. 2004; 44: 809–21. doi: 10.1016/j.neuron.2004.11.019 PMID: 15572112

42. Shazman S, Celniker G, Haber O, Glaser F, Mandel-Gutfreund Y. Patch Finder Plus (PFplus): a web server for extracting and displaying positive electrostatic patches on protein surfaces. Nucleic Acids Res. 2007; 35: W526–30. doi: 10.1093/nar/gkm401 PMID: 17537808

43. Scrima A, Konícková R, Czyzewski BK, Kawasaki Y, Jeffrey PD, Groisman R, et al. Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. Cell. 2008; 135: 1213–23. doi: 10.1016/j.cell.2008.10.045 PMID: 19109893

44. Verspurten J, Gevaert K, Declercq W, Vandenabeele P. SitePredicting the cleavage of proteinase substrates. Trends Biochem Sci. 2009; 34: 319–23. doi: 10.1016/j.tibs.2009.04.001 PMID: 19546006

45. Chevaillier P. Pest sequences in nuclear proteins. Int J Biochem. 1993; 25: 479–482. doi: 10.1016/0020-711X(93)90653-V PMID: 8096823

46. Smith TF, Gaitatzes C, Saxena K, Neer EJ. The WD repeat: a common architecture for diverse functions. Trends Biochem Sci. 1999; 24: 181–5. Available: http://www.ncbi.nlm.nih.gov/pubmed/10322433 PMID: 10322433

47. Xu C, Min J. Structure and function of WD40 domain proteins. Protein Cell. 2011; 2: 202–14. doi: 10.1007/s13238-011-1018-1 PMID: 21468892